

# WEIGHTED BERGMAN SPACES: SHIFT-INVARIANT SUBSPACES AND INPUT/STATE/OUTPUT LINEAR SYSTEMS

JOSEPH A. BALL AND VLADIMIR BOLOTNIKOV

**ABSTRACT.** It is well known that subspaces of the Hardy space over the unit disk which are invariant under the backward shift occur as the image of an observability operator associated with a discrete-time linear system with stable state-dynamics, as well as the functional-model space for a Hilbert space contraction operator, while forward shift-invariant subspaces have a representation in terms of an inner function. We discuss several variants of these statements in the context of weighted Bergman spaces on the unit disk.

## 1. INTRODUCTION

Let  $H^2$  be the standard Hardy space over the open unit disk

$$H^2 = \{f(z) = \sum_{n=0}^{\infty} f_n z^n : \sum_{n=0}^{\infty} |f_n|^2 < \infty\}$$

and we let  $S: f(z) \mapsto zf(z)$  be the shift operator on  $H^2$ . The classical theorem of Beurling [11] asserts that any  $S$ -invariant subspace  $\mathcal{M}$  of  $H^2$  has the form  $\mathcal{M} = \theta \cdot H^2$  where  $\theta$  is an inner function (analytic on the disk with boundary-value function on the unit circle having modulus 1 almost everywhere). The result was extended to the vector-valued case by Lax [22] (for the finite-dimensional case) and by Halmos [18] (for the general case). The main thrust of Beurling's paper was the development of a theory of inner-outer factorization for  $H^2$  and  $H^\infty$ -functions; this theory was then used to arrive at the famous characterization of invariant subspaces  $\mathcal{M} = \theta \cdot H^2$ .

There are now a number of more operator-theoretic and/or system-theoretic proofs which better handle the vector-valued case. For  $\mathcal{Y}$  a Hilbert space, we write  $H^2(\mathcal{Y})$  for the Hardy space  $H^2(\mathcal{Y}) = H^2 \otimes \mathcal{Y}$  of  $\mathcal{Y}$ -valued functions and we let  $\mathcal{M}$  be a shift-invariant subspace of  $H^2(\mathcal{Y})$ . We single out four distinct approaches toward what we shall call simply the Beurling-Lax theorem for the vector-valued case.

(1) Find a Hilbert space  $\mathcal{X}$  (playing the role of a *state space* from the system theory point of view) and construct operators  $C: \mathcal{X} \rightarrow \mathcal{Y}$  and  $A: \mathcal{X} \rightarrow \mathcal{X}$  so that  $\begin{bmatrix} A \\ C \end{bmatrix}$  is isometric ( $A^*A + C^*C = I_{\mathcal{X}}$ ) and so that  $\mathcal{M}^\perp$  has the representation

$$\mathcal{M}^\perp = \{C(I - zA)^{-1}x : x \in \mathcal{X}\}. \quad (1.1)$$

One such choice is  $\mathcal{X} = \mathcal{M}^\perp$  with  $C = E|_{\mathcal{M}^\perp}$  (where  $E: f \mapsto f(0)$  is evaluation at zero and  $A = S^*|_{\mathcal{M}^\perp}$ ). We note that  $H^2(\mathcal{Y})$  is a reproducing kernel Hilbert space with reproducing kernel given by the Szegő kernel  $\frac{I_{\mathcal{Y}}}{1-z\bar{\zeta}}$ . Then  $\mathcal{M}^\perp \subset H^2(\mathcal{Y})$ , as a

---

1991 *Mathematics Subject Classification.* 47A48, 47A57.

*Key words and phrases.* Operator-valued functions, Bergman space, Beurling-Lax representations, transfer-function realization.

subspace of  $H^2(\mathcal{Y})$ , is also a reproducing kernel Hilbert space in its own right. One can identify the reproducing kernel for  $\mathcal{M}^\perp$  explicitly as

$$\mathcal{M}^\perp = \mathcal{H}(K_{C,A}) \quad \text{where} \quad K_{C,A} = C(I - zA)^{-1}(I - \bar{\zeta}A^*)^{-1}C^*.$$

It then follows that  $\mathcal{M}$  has Beurling-Lax-Halmos representation  $\mathcal{M} = \Theta \cdot H^2(\mathcal{U})$  where  $\Theta$  is any solution of the kernel factorization problem

$$\frac{\Theta(z)\Theta(\zeta)^*}{1 - z\bar{\zeta}} = \frac{I_{\mathcal{Y}}}{1 - z\bar{\zeta}} - K_{C,A}(z, \zeta) \quad (1.2)$$

from which it follows that the reproducing kernel  $k_{\mathcal{M}}(z, \zeta)$  for the subspace  $\mathcal{M}$  is

$$k_{\mathcal{M}}(z, \zeta) = \frac{\Theta(z)\Theta(\zeta)^*}{1 - z\bar{\zeta}}. \quad (1.3)$$

One way to construct such a  $\Theta$  is as

$$\Theta(z) = D + zC(I - zA)^{-1}B \quad (1.4)$$

where  $\begin{bmatrix} B \\ D \end{bmatrix}$  solves the Cholesky factorization problem

$$\begin{bmatrix} B \\ D \end{bmatrix} \begin{bmatrix} B^* & D^* \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} A \\ C \end{bmatrix} \begin{bmatrix} A^* & C^* \end{bmatrix}.$$

We refer to this approach as the *de Branges-Rovnyak-Potapov approach*, as the first part (the identification of  $\mathcal{M}^\perp$  as  $\mathcal{H}(K_{C,A})$ ) is prominent in the work of de Branges-Rovnyak [12, 13] while the second step (factorization of the kernel (1.2)) is prominent in the Potapov-school approach to interpolation (see [17]).

**(2)** Set  $A = P_{\mathcal{M}}S^*|_{\mathcal{M}}$  and  $C = P_{\mathcal{E}}|_{\mathcal{M}}$ ,  $\mathcal{U} = \mathcal{M} \ominus S\mathcal{M}$  and consider the operator  $\mathcal{O}_{C,A}: \mathcal{M} \subset H^2(\mathcal{Y}) \rightarrow H^2(\mathcal{U})$  given by

$$\mathcal{O}_{C,A}: x \mapsto C(I - zA)^{-1}x.$$

Then the inner representer  $\Theta$  for  $\mathcal{M}$  can be obtained as

$$M_\Theta = (\mathcal{O}_{C,A})^*: H^2(\mathcal{U}) \rightarrow H^2(\mathcal{Y}).$$

This approach works more generally for  $S$ -invariant subspaces of  $H^2(\mathcal{Y})$  contractively included in  $H^2(\mathcal{Y})$ . We refer to this approach as the *Rosenblum-Rovnyak approach*, as this is amounts to the proof of the Beurling-Lax Theorem in the book [30].

**(3)** Note that  $\mathcal{M}$  decomposes as  $\mathcal{M} = \bigoplus_{k \geq 0} (S^k\mathcal{M} \ominus S^{k+1}\mathcal{M})$ . It turns out that the reproducing kernel for the subspace  $S^k\mathcal{M} \ominus S^{k+1}\mathcal{M}$  has the form

$$k_{S^k\mathcal{M} \ominus S^{k+1}\mathcal{M}}(z, \zeta) = z^k \bar{\zeta}^k \Theta(z)\Theta(\zeta)^* \quad (1.5)$$

with  $\Theta(z)$  independent of  $k$ . Thus the operator of multiplication by  $z^k\Theta(z)$  maps the coefficient space  $\mathcal{U}$  isometrically onto  $S^k\mathcal{M} \ominus S^{k+1}\mathcal{M}$  and we recover the kernel function for  $\mathcal{M}$  to be

$$k_{\mathcal{M}}(z, \zeta) = \sum_{k=0}^{\infty} z^k \bar{\zeta}^k \Theta(z)\Theta(\zeta)^*$$

in agreement with (1.3). It now follows that  $\Theta$  is a Beurling-Lax representer for  $\mathcal{M}$ . We refer to this as the *iterated Halmos wandering-subspace approach* since the construction of the inner representer via looking at the wandering subspace  $\mathcal{M} \ominus S\mathcal{M}$  is the main idea of the proof for the abstract shift-operator setting in [18].

(4) Construct  $\Theta$  so that  $M_\Theta \mathcal{U} = \mathcal{M} \ominus S\mathcal{M}$ , or  $k_{\mathcal{M} \ominus S\mathcal{M}} = \Theta(z)\Theta(\zeta)^*$  (i.e.,  $\Theta$  is the same as in Approach 3). In this shift-operator context, then the multiplication operator  $M_\Theta$  extends to be an isometry from  $H^2(\mathcal{U})$  into  $H^2(\mathcal{Y})$  with  $M_\Theta H^2(\mathcal{U}) = \mathcal{M}$  and once again  $k_{\mathcal{M}}(z, \zeta) = \Theta(z) \left( \frac{I_{\mathcal{U}}}{1-z\bar{\zeta}} \right) \Theta(\zeta)^*$ , so once again  $\Theta$  is the Beurling-Lax representer for  $\mathcal{M}$ . We refer to this as the *uniterated Halmos wandering subspace approach*.

In all four approaches, one ultimately arrives at the realization formula (1.4) for the inner function  $\Theta$  where the system matrix (also called operator colligation)  $\mathbf{U} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$  is unitary. This has the interpretation that the associated input/state/output linear system

$$\Sigma_{\mathbf{U}} : \begin{cases} x(n+1) = Ax(n) + Bu(n) \\ y(n) = Cx(n) + Du(n) \end{cases} \quad (1.6)$$

is *conservative*: the energy stored by the state at time  $n$  ( $\|x(n+1)\|^2 - \|x(n)\|^2$ ) is exactly compensated by the net energy put into the system from the outside environment ( $\|u(n)\|^2 - \|y(n)\|^2$ ), with a similar property for the adjoint system. We note that application of the  $Z$ -transform

$$\{x(n)\}_{n \in \mathbb{Z}_+} \mapsto \widehat{x}(z) := \sum_{n=0}^{\infty} x(n)z^n$$

converts the system equations (1.6) to

$$\begin{aligned} z^{-1}(\widehat{x}(z) - x(0)) &= A\widehat{x}(z) + B\widehat{u}(z) \\ \widehat{y}(z) &= C\widehat{x}(z) + D\widehat{u}(z) \end{aligned} \quad (1.7)$$

which can then be solved for  $\widehat{x}(z)$  and  $\widehat{y}(z)$ :

$$\begin{aligned} \widehat{x}(z) &= (I - zA)^{-1}x(0) + z(I - zA)^{-1}B\widehat{u}(z) \\ \widehat{y}(z) &= (\mathcal{O}_{C,A}x(0))(z) + \Theta_{\mathbf{U}}(z) \cdot \widehat{u}(z) \end{aligned}$$

where

$$\mathcal{O}_{C,A} : x \mapsto C(I - zA)^{-1}x \quad (1.8)$$

is the *observability operator* for the system  $\Sigma_{\mathbf{U}}$  (generating the  $Z$ -transform  $\widehat{y}(z)$  of the output signal from the initial state  $x(0) = x$  when the input signal is taken to be zero), and where

$$\Theta_{\mathbf{U}}(z) = D + zC(I - zA)^{-1}B \quad (1.9)$$

is the *transfer function* of the system  $\Sigma_{\mathbf{U}}$  (having the property that multiplication by  $\Theta_{\mathbf{U}}$  on the  $Z$ -transform  $\widehat{u}(z)$  of the input signal generates the  $Z$ -transform  $\widehat{y}(z)$  of the output when the initial state  $x(0)$  is taken to be zero). We note that the observability operator  $\mathcal{O}_{C,A}$  appears in the representation (1.1) for  $\mathcal{M}^\perp$  and that the formula (1.4) exhibits the Beurling-Lax representer for  $\mathcal{M}$  as the transfer function for the conservative i/s/o linear system  $\Sigma_{\mathbf{U}}$ . Many of these ideas connecting operator-valued  $H^\infty$ -functions with systems theory ideas can be found already in the survey paper of Ball-Cohen [10].

It was only much later that researchers began looking for analogues of Beurling-Lax representations for shift-invariant subspaces of weighted Bergman spaces (see [19, 15] and the references there for the scalar-valued case). For a Hilbert space  $\mathcal{Y}$  and an integer  $n \geq 1$ , we denote by  $\mathcal{A}_n(\mathcal{Y})$  the reproducing kernel Hilbert space with reproducing kernel

$$\mathbf{k}_n(z, \zeta) = \frac{1}{(1 - z\bar{\zeta})^n} I_{\mathcal{Y}} \quad (1.10)$$

which is clearly positive on  $\mathbb{D} \times \mathbb{D}$ . The space  $\mathcal{A}_n(\mathcal{Y})$  can be alternatively characterized as the Hilbert space of  $\mathcal{Y}$ -valued functions  $f$  analytic in the open unit disk  $\mathbb{D}$  and with finite norm  $\|f\|_{\mathcal{A}_n(\mathcal{Y})}$ :

$$\mathcal{A}_n(\mathcal{Y}) = \left\{ f(z) = \sum_{j \geq 0} f_j z^j : \|f\|_{\mathcal{A}_n(\mathcal{Y})}^2 := \sum_{j \geq 0} \mu_{n,j} \cdot \|f_j\|_{\mathcal{Y}}^2 < \infty \right\} \quad (1.11)$$

where the weights  $\mu_{n,k}$ 's are defined by

$$\mu_{n,j} := \frac{1}{\binom{j+n-1}{j}} = \frac{j!(n-1)!}{(j+n-1)!}. \quad (1.12)$$

We let  $S_n$  denote the Bergman shift operator of weight-index  $n$  acting on  $\mathcal{A}_n(\mathcal{Y})$  given by multiplication by the coordinate functions:  $S_n : f(z) \mapsto zf(z)$  for  $f \in \mathcal{A}_n(\mathcal{Y})$ . The space  $A_1(\mathcal{Y})$  is therefore the standard Hardy space  $H^2(\mathcal{Y})$  of the unit disk with  $S_1 = S$  equal to the standard Hardy-space shift operator, and for  $n \geq 2$  the space  $\mathcal{A}_n(\mathcal{Y})$  is the standard weighted Bergman space.

Even for the case  $\mathcal{Y} = \mathbb{C}$ , one can no longer use Blaschke products to collect the zeros of a Bergman-space function since the zero set of a Bergman space function need not satisfy the Blaschke growth condition. A major advance came with the work of Hedenmalm [20] who constructed Bergman-inner functions, i.e., a function  $\theta$  so that  $\theta\mathbb{C} = \mathcal{A}_2 \ominus S_2\mathcal{A}_2$  as the solution of an extremal problem. Such a  $\theta$  has the *contractive divisor property*  $\|\theta^{-1}f\|_{\mathcal{A}_2} \leq \|f\|_{\mathcal{A}_2}$ , thereby improving earlier results of Horowitz [21]. Since the work of Apostol-Bercovici-Foias-Pearcy [5] it is known that shift-invariant subspaces  $\mathcal{M} \subset \mathcal{A}_2$  can have arbitrary index  $\text{ind } \mathcal{M} := \dim(\mathcal{M} \ominus S_2\mathcal{M})$ . Nevertheless, the seminal work of Aleman-Richter-Sundberg [2] with later extensions by Shimorin [31] showed that in all cases we recover  $\mathcal{M}$  as  $\mathcal{M} = \overline{\text{span}}_{k \geq 0} S_n^k(\mathcal{M} \ominus S_n\mathcal{M})$  (a partial analogue of the representation of  $\mathcal{M}$  in approach 4 above), at least for the cases  $n = 2, 3$ . More recently, Olofsson [27, 28, 29] initiated the study of operator-valued Bergman-inner functions for standard weighted Bergman spaces as an object of independent interest. In particular he suggested a certain time-invariant input/state/output linear system of higher order (past values of the states and of the inputs enter the state-update equations rather than just the current state and input values) as the time-domain explanation for the input-output map corresponding to multiplication by the Bergman inner function.

The purpose of the present paper is to further enhance the theory of Beurling-Lax representations and to sharpen the connections with the theory of input/state/output linear systems for the Bergman-space setting. In particular, we show that each of the four approaches to Beurling-Lax representations sketched above (which blur into each other in the classical case) actually lead to four distinct kinds of theories for the Bergman-space setting. Our results for the first three approaches are new for the Bergman-space setting, while the fourth approach is most closely connected with the notion of Bergman inner function as appearing in the work of Hedenmalm, Olofsson, and Shimorin. Specifically, for  $\mathcal{M} \subset \mathcal{A}_n(\mathcal{Y})$ , the first approach leads to the representation for  $\mathcal{M}$  as

$$\mathcal{M} = M_F \left( \bigoplus_{j=1}^n \mathcal{A}_j(\mathcal{U}_j) \right), \quad \text{where} \quad F(z) = [F_1(z) \quad \cdots \quad F_n(z)]$$

and where  $F_j$  is a bounded multiplier from  $\mathcal{A}_j(\mathcal{U}_j)$  into  $\mathcal{A}_n(\mathcal{Y})$  for each  $j = 1, \dots, n$  with the additional property that the multiplication operator  $[M_{F_1} \ \dots \ M_{F_n}]$  is a partial isometry from  $\bigoplus_{j=1}^n \mathcal{A}_j(\mathcal{U}_j)$  into  $\mathcal{A}_n(\mathcal{Y})$  (see Theorem 7.1). The second approach leads to the identification of certain conditions when one has a representation  $\mathcal{M} = M_{F_n} \mathcal{A}_n(\mathcal{U}_n)$  (i.e., one can take  $F_1, \dots, F_{n-1}$  all equal to zero in the previous representation) and applies to the more general situation where  $\mathcal{M}$  has its own norm and is contractively included in  $\mathcal{A}_n(\mathcal{Y})$ . The third approach leads to the construction of a *Bergman inner family*  $\Theta_0, \Theta_1, \Theta_2, \dots$  of multipliers, where  $z^k \Theta_k(z)$  maps the coefficient space  $\mathcal{U}_k$  isometrically into  $\mathcal{A}_n(\mathcal{Y})$  so that  $[M_{\Theta_0} \ M_{\Theta_1} \ M_{\Theta_2} \ \dots]$  maps the time-varying Hardy space  $\bigoplus_{k=0}^{\infty} z^k \mathcal{U}_k$  isometrically onto  $\mathcal{M} \subset \mathcal{A}_n(\mathcal{Y})$ . This perhaps is the most compelling new Beurling-Lax representation for the Bergman-space setting and has the most striking connections with systems theory. Namely, the multiplication operator  $[M_{\Theta_0} \ M_{\Theta_1} \ M_{\Theta_2} \ \dots]$  can be identified as the  $Z$ -transform of the input-output map of a certain time-varying linear system (see formula (3.2) below) having certain additional metric properties (see (6.7), (6.22) and Section 8). When one specializes to the classical case  $n = 1$ , this time-varying linear system collapses to the time-invariant linear system (1.6) with transfer function equal to the Beurling-Lax representer  $\Theta(z)$  in the classical case.

To make these system theory connections precise, we delve into the structure of observability operators and observability gramians, operator resolvents, and transfer functions more general than in the classical case. Namely, observability operators of the form

$$\mathcal{O}_{n,C,A}: x \mapsto C(I - zA)^{-n}x = \sum_{j=0}^{\infty} \binom{j+n-1}{j} (A^j x) z^j \quad (1.13)$$

as well as a  $k$ -shifted version

$$\mathfrak{D}_{n,k,C,A}: x \mapsto \sum_{j=0}^{\infty} \binom{j+k+n-1}{j+k} (A^j x) z^j,$$

as well as functions  $\Theta$  having a realization of the form

$$\Theta(z) = D + z\mathfrak{D}_{n,k,C,A}(z)B \quad (1.14)$$

come up. Indeed such formulas appear already in the work of Olofsson [29], at least for the cases  $k = 0, 1$ . We show that  $\mathcal{O}_{n,C,A}$  (1.13) arises as the observability operator for the time-varying system (3.2) and that operator-valued functions of the form  $\Theta$  (1.14) arise naturally in the construction of kernel functions for subspaces of the form  $S_n^k \mathcal{M} \ominus S_n^{k+1} \mathcal{M}$  (where  $\mathcal{M}$  is an  $S_n$ -invariant subspace of  $\mathcal{A}_n(\mathcal{Y})$ ).

As preliminaries to the exposition of these ideas, in Section 2 we develop a calculus of shifted power geometric series as preparation for their use in the corresponding operator functional calculus. In Section 3 we present the form of the time-varying system (3.2) and develop its behavior under the  $Z$ -transform for the general setting where no metric constraints are imposed. Section 4 extends standard stability notions to the Bergman-setting where the system has the form (3.2). Section 5 identifies the reproducing-kernel structure on ranges of observability operators; a key result is Theorem 5.6 which identifies the reproducing kernel for a subspace of  $\mathcal{A}_n(\mathcal{Y})$  of the form  $S_n^k \text{Ran } \mathfrak{D}_{n,k,C,A}$ . The next Section 6 develops the metric properties for the system operators in the system (3.2) which will be needed for the

Bergman-space Beurling-Lax representations to come. Finally Section 7 develops our Beurling-Lax representation theorems for the Bergman-space setting (all four approaches) while the concluding Section 8 makes the connections with the linear system (3.2) (with the additional metric constraints imposed) precise.

## 2. POWER-SERIES REPRESENTATIONS FOR GENERALIZED GEOMETRIC SERIES

We start with recording an assortment of power series expansions which will play a key role in the sequel. We let  $S_1^*$  be the standard backward shift operator acting on formal power series according to

$$S_1^*: \sum_{j=0}^{\infty} a_j z^j \mapsto \sum_{j=0}^{\infty} a_{j+1} z^j \text{ or, equivalently, } S_1^*: f(z) \mapsto \frac{f(z) - f(0)}{z}. \quad (2.1)$$

Let us introduce the notation

$$R_n(z) = (1-z)^{-n} \quad \text{and} \quad R_{n,k}(z) = (S_1^*)^k (1-z)^{-n}, \quad (2.2)$$

so that  $R_{n,0}(z) = R_n(z)$ , and record the power series expansions

$$R_n(z) = \sum_{j=0}^{\infty} \binom{n+j-1}{j} \cdot z^j, \quad R_{n,k}(z) = \sum_{j=0}^{\infty} \binom{n+j+k-1}{j+k} \cdot z^j. \quad (2.3)$$

The first representation follows from successive term-by-term differentiation of the geometric series  $(1-z)^{-1} = \sum_{j \geq 0} z^j$  whereas the second follows from the first and the definition (2.1) of  $S_1^*$ . It is seen from (2.3) that for the special case  $n = 1$ ,

$$R_{1,k}(z) = R_1(z) \quad \text{for all } k = 0, 1, \dots \quad (2.4)$$

The formula

$$R_{n,k}(z) = \sum_{\ell=1}^n \binom{\ell+k-2}{\ell-1} R_{n-\ell+1}(z) \quad \text{for } k \geq 1 \quad (2.5)$$

follows from (2.3) and the Chu-Vandermonde identity for binomial coefficients (see e.g. [23, page 50])

$$\binom{n+j+k-1}{j+k} = \sum_{\ell=1}^n \binom{\ell+k-2}{\ell-1} \cdot \binom{n+j-\ell}{j}, \quad (2.6)$$

according to which indeed

$$\begin{aligned} R_{n,k}(z) &= \sum_{j=0}^{\infty} \sum_{\ell=1}^n \binom{\ell+k-2}{\ell-1} \binom{n+j-\ell}{j} z^j = \sum_{\ell=1}^n \binom{\ell+k-2}{\ell-1} \cdot \left( \sum_{j=0}^{\infty} \binom{n+j-\ell}{j} z^j \right) \\ &= \sum_{\ell=1}^n \binom{\ell+k-2}{\ell-1} \cdot R_{n-\ell+1}(z) \end{aligned}$$

where we made use of (2.3) again in the last step. An easy corollary of definitions (2.2) is the following.

**Lemma 2.1.** *The functions  $R_{n,k}(z)$  given by (2.2) satisfy the recursion*

$$R_{n,k}(z) = \binom{n+k-1}{k} + z R_{n,k+1}(z). \quad (2.7)$$

*Proof.* Since  $R_{n,k+1}(z) = (S_1^*)^{k+1} (1-z)^{-n} = S_1^* ((S_1^*)^k (1-z)^{-n}) = S_1^* R_{n,k}(z)$ , we have

$$R_{n,k+1}(z) = S_1^* R_{n,k}(z) = [R_{n,k}(z) - R_{n,k}(0)]/z.$$

From the expansion for  $R_{n,k}(z)$  in (2.3) we see that  $R_{n,k}(0) = \binom{n+k-1}{k}$  and the formula (2.7) follows.  $\square$

The following analogue of the well-known formula  $\sum_{j=0}^N z^j = \frac{1-z^{N+1}}{1-z}$  for the sum of the truncated geometric series will be useful in the sequel.

**Proposition 2.2.** *The truncation of the infinite series representation for  $R_n(z)$  in (2.3) has the explicit summation formula:*

$$\sum_{j=0}^N \binom{n+j-1}{j} z^j = \frac{1}{(1-z)^n} - \sum_{j=1}^n \binom{N+n}{N+j} \frac{z^{N+j}}{(1-z)^j}. \quad (2.8)$$

*Proof.* The following computation

$$\begin{aligned} \sum_{j=0}^N \binom{n+j-1}{j} z^j &= \frac{1}{(n-1)!} \cdot \frac{d^{n-1}}{dz^{n-1}} \left( \frac{1-z^{N+n}}{1-z} \right) \\ &= \frac{1-z^{N+n}}{(1-z)^n} - \frac{1}{(n-1)!} \cdot \sum_{j=0}^{n-2} \binom{n-1}{j} \frac{j!(N+k)!}{(N+j+1)!} \frac{z^{N+j+1}}{(1-z)^{j+1}} \\ &\quad (\text{by the Leibnitz rule for the derivative of a product}) \\ &= \frac{1-z^{N+k}}{(1-z)^k} - \sum_{j=0}^{n-2} \binom{N+n}{N+j+1} \frac{z^{N+j+1}}{(1-z)^{j+1}} \\ &= \frac{1}{(1-z)^n} - \sum_{j=1}^n \binom{N+n}{N+j} \frac{z^{N+j}}{(1-z)^j} \quad (\text{by the shift } j \mapsto j+1) \end{aligned}$$

verifies the result.  $\square$

Proposition 2.2 enables us to express the shifted  $n$ -th-power geometric series  $R_{n,k}(z)$  in terms of unshifted  $n$ -th-power geometric series  $R_n(z)$  as follows.

**Proposition 2.3.** *The shifted  $n$ -resolvent  $R_{n,k}(z)$  is recovered from the unshifted  $n$ -resolvent  $R_n(z)$  according to the formula*

$$R_{n,k}(z) = R_n(z) \cdot \sum_{\kappa=0}^{n-1} \left( \sum_{j=0}^{\kappa} (-1)^j \binom{k+n-1}{k} \binom{n-1-\kappa+j}{j} \right) z^\kappa \quad (2.9)$$

*Proof.* By definitions (2.2),

$$\begin{aligned} R_{n,k}(z) &= z^{-k} \cdot \left( R_n(z) - \sum_{j=0}^{k-1} \binom{n+j-1}{j} z^j \right) \\ &= z^{-k} R_n(z) \cdot \left( 1 - (1-z)^n \cdot \sum_{j=0}^{k-1} \binom{n+j-1}{j} z^j \right). \end{aligned} \quad (2.10)$$

Letting  $N = k - 1$  in (2.8) and combining this formula with (2.10) gives

$$\begin{aligned} R_{n,k}(z) &= R_n(z) \cdot \sum_{j=0}^{n-1} \binom{k+n-1}{k} z^j (1-z)^{n-j-1} \\ &= R_n(z) \cdot \sum_{j=0}^{n-1} \sum_{\ell=0}^{n-j-1} \binom{k+n-1}{k} \binom{n-j-1}{\ell} (-1)^\ell z^{j+\ell} \\ &= R_n(z) \cdot \sum_{\kappa=0}^{n-1} \left( \sum_{j=0}^{\kappa} \binom{k+n-1}{k} \binom{n-1-j}{\kappa-j} (-1)^{\kappa-j} \right) z^\kappa \\ &= R_n(z) \cdot \sum_{\kappa=0}^{n-1} \left( \sum_{j=0}^{\kappa} \binom{k+n-1}{k} \binom{n-1-\kappa+j}{j} (-1)^j \right) z^\kappa \end{aligned}$$

and (2.9) follows.  $\square$

As a consequence of identities (2.3), (2.5), (2.7), (2.9), we have the following identities for any Hilbert-space operator  $A$  having spectral radius less than one:

$$R_n(zA) = (I - zA)^{-n} = \sum_{j=0}^{\infty} \binom{n+j-1}{j} A^j z^j, \quad (2.11)$$

$$R_{n,k}(zA) = \sum_{j=0}^{\infty} \binom{n+j+k-1}{j+k} A^j z^j, \quad (2.12)$$

$$R_{n,k}(zA) = \sum_{\ell=1}^n \binom{\ell+k-2}{\ell-1} R_{n-\ell+1}(zA) \quad \text{for } k \geq 1, \quad (2.13)$$

$$R_{n,k}(zA) = \binom{n+k-1}{k} I_{\mathcal{X}} + zAR_{n,k+1}(zA), \quad (2.14)$$

$$R_{n,k}(zA) = R_n(zA) \cdot \sum_{\kappa=0}^{n-1} \left( \sum_{j=0}^{\kappa} (-1)^j \binom{k+n-1}{k} \binom{n-1-\kappa+j}{j} \right) z^\kappa A^\kappa. \quad (2.15)$$

### 3. SYSTEM THEORY MOTIVATION

We have seen in the Introduction (see formulas (1.6)–(1.9)) how a formula of the type

$$\Theta_{\Sigma_1}(z) = D + zC(I - zA)^{-1}B = D + zCR_1(zA)B. \quad (3.1)$$

arises as the transfer function of a discrete-time time-invariant input/state/output linear system of the form (1.6). Furthermore, imposition of the condition that the colligation matrix  $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$  be contractive (respectively, unitary) together with an additional stability condition on the state-update operator  $A$  leads to the transfer function  $\Theta$  being in the Schur class, i.e., having contractive values on the unit disk (respectively, being inner, i.e., contractive on the unit disk with unitary nontangential boundary values on the unit circle almost everywhere).

In the quest for a parallel system-theoretic interpretation for Bergman inner functions on the unit disk, Olofsson [27, 29] has shown that realization formulas of the type  $\Theta(z) = D + zCR_n(zA)B$  involving higher-order resolvents  $R_n(zA)$  arise naturally and he also associated a certain higher-order time-invariant linear system having such a  $\Theta(z)$  as transfer function. We shall see below that functions of the

form  $\Theta(z) = D + zCR_{n,k}(zA)B$  involving shifted higher-order resolvents  $R_{n,k}(zA)$  also arise naturally. To obtain a system-theoretic connection for functions involving such shifted higher-order resolvents  $R_{n,k}(zA)$ , we propose to consider the following discrete-time time-varying linear system:

$$\Sigma_n \left( \left\{ \begin{bmatrix} A & B_j \\ C & D_j \end{bmatrix} \right\}_{j \in \mathbb{Z}_+} \right) : \begin{cases} x(j+1) = \frac{j+n}{j+1} \cdot Ax(j) + \binom{j+n}{j+1} \cdot B_j u(j), \\ y(j) = Cx(j) + \binom{j+n-1}{j} \cdot D_j u(j) \end{cases} \quad (3.2)$$

where

$$A \in \mathcal{L}(\mathcal{X}), \quad C \in \mathcal{L}(\mathcal{X}, \mathcal{Y}), \quad B_k \in \mathcal{L}(\mathcal{U}_k, \mathcal{X}), \quad D_k \in \mathcal{L}(\mathcal{U}_k, \mathcal{Y})$$

are given bounded linear operators acting between given Hilbert spaces  $\mathcal{X}$ ,  $\mathcal{Y}$  and  $\mathcal{U}_k$  ( $k \geq 0$ ). We note that the case where  $n = 1$  and where the operators  $B_k = B$  and  $D_k = D$  are taken independent of the time parameter  $k \in \mathbb{Z}_+$  reduces to the classical time-invariant case given by (1.6).

If we let the system evolve on the nonnegative integers  $j \in \mathbb{Z}_+$ , then the whole trajectory  $\{u(j), x(j), y(j)\}_{j \in \mathbb{Z}_+}$  is determined from the input signal  $\{u(j)\}_{j \in \mathbb{Z}_+}$  and the initial state  $x(0)$  according to the formulas

$$x(j) = \binom{n+j-1}{j} \cdot \left( A^j x(0) + \sum_{\ell=0}^{j-1} A^{j-\ell-1} B_\ell u(\ell) \right), \quad (3.3)$$

$$y(j) = \binom{n+j-1}{j} \cdot \left( CA^j x(0) + \sum_{\ell=0}^{j-1} CA^{j-\ell-1} B_\ell u(\ell) + D_j u(j) \right). \quad (3.4)$$

Formula (3.3) is established by simple induction arguments, while (3.4) is obtained by substituting (3.3) into the second equation in (3.2).

To write the  $Z$ -transformed version of the system-trajectory formula (3.3), we multiply both sides of (3.3) by  $z^j$  and sum over  $j \geq 0$  to get, on account of (2.11),

$$\begin{aligned} \widehat{x}(z) &= \sum_{j=0}^{\infty} x(j) z^j \\ &= \left( \sum_{j=0}^{\infty} \binom{j+n-1}{j} A^j z^j \right) x(0) + \sum_{k=1}^{\infty} \left( \sum_{j=k}^{\infty} \binom{j+n-1}{j} A^{j-k} z^j \right) B_{k-1} u(k-1) \\ &= (I - zA)^{-n} x(0) + \sum_{k=1}^{\infty} z^k \left( \sum_{j=0}^{\infty} \binom{j+k+n-1}{j+k} A^j z^j \right) B_{k-1} u(k-1) \\ &= (I - zA)^{-n} x(0) + \sum_{k=1}^{\infty} z^k R_{n,k}(zA) B_{k-1} u(k-1) \\ &= (I - zA)^{-n} x(0) + \sum_{k=0}^{\infty} z^{k+1} R_{n,k+1}(zA) B_k u(k). \end{aligned} \quad (3.5)$$

The same procedure applied to (3.4) gives

$$\begin{aligned}\widehat{y}(z) &= \sum_{k=0}^{\infty} y(k) z^k \\ &= C(I - zA)^{-n} x(0) + \sum_{k=0}^{\infty} z^k \left( \binom{k+n-1}{k} D_k + zCR_{n,k+1}(zA)B_k \right) u(k) \\ &= \mathcal{O}_{n,C,A}x(0) + \sum_{k=0}^{\infty} z^k \Theta_{n,k}(z) u(k),\end{aligned}\tag{3.6}$$

where

$$\mathcal{O}_{n,C,A}: x \mapsto \sum_{j=0}^{\infty} \left( \binom{n+j-1}{j} CA^j x \right) z^j = C(I - zA)^{-n} x\tag{3.7}$$

is the  $n$ -observability operator and where

$$\Theta_{n,k}(z) = \binom{k+n-1}{k} D_k + zCR_{n,k+1}(zA)B_k \quad (k = 0, 1, \dots)\tag{3.8}$$

is the family of transfer functions. We note that the transfer function  $\Theta_{n,k}(z)$  encodes the result of a pulse input-vector  $u$  being applied at time  $j = k$ :

$$\widehat{y}(z) = \Theta_{n,k}(z) \cdot z^k u \quad \text{if } x(0) = 0 \quad \text{and} \quad u(j) = \delta_{j,k} u.$$

In fact the functions  $\Theta_{n,k}(z)$  could have been derived in this way and then one could arrive at input-output relation (3.6) via superposition of all these time- $k$  impulse responses. Note also that formula (3.8) for the classical time-invariant case ( $n = 1$  and  $B_k, D_k$  independent of  $k$ ) reduces to the formula (3.1) for the classical transfer function  $\Theta_{\Sigma_1}$  due to the identity (2.4).

For the application to Bergman inner functions, it is natural to impose some additional metric constraints on the colligation matrices  $\begin{bmatrix} A & B_k \\ C & D_k \end{bmatrix}$ ; these are discussed in Section 6 below.

#### 4. OBSERVABILITY OPERATORS AND GRAMIANS, STEIN EQUALITIES AND INEQUALITIES

Formula (3.7) associates with any *output pair*  $(C, A)$  (i.e.,  $C \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$  and  $A \in \mathcal{L}(\mathcal{X})$ ) the  $n$ -observability operator  $\mathcal{O}_{n,C,A}$ . In case  $\mathcal{O}_{n,C,A}$  is bounded as an operator from  $\mathcal{X}$  into  $\mathcal{A}_n(\mathcal{Y})$ , we say that the pair  $(C, A)$  is  *$n$ -output-stable*. If  $(C, A)$  is  $n$ -output stable, then the  *$n$ -observability gramian*

$$\mathcal{G}_{n,C,A} := (\mathcal{O}_{n,C,A})^* \mathcal{O}_{n,C,A}\tag{4.1}$$

is bounded on  $\mathcal{X}$  and can be represented via the series

$$\mathcal{G}_{n,C,A} = \sum_{k=0}^{\infty} \binom{k+n-1}{k} A^{*k} C^* C A^k\tag{4.2}$$

converging in the strong operator topology. As suggested by the Agler hereditary functional calculus as formulated by Ambrozie-Engliš-Müller [4], we introduce the operator

$$B_A: X \mapsto A^* X A\tag{4.3}$$

mapping  $\mathcal{L}(\mathcal{X})$  into itself, and then view  $\mathcal{G}_{n,C,A}$  (at least formally) as being given by

$$\mathcal{G}_{n,C,A} = (I - B_A)^{-n} [C^* C].\tag{4.4}$$

If  $\|B_A\| < 1$ , (4.4) is precise; in general one can make this precise by interpreting (4.4) in the form

$$\mathcal{G}_{n,C,A} = \lim_{r \uparrow 1} (I - rB_A)^{-n} [C^* C]. \quad (4.5)$$

Either of the formulas (4.4) and (4.5) suggests that we define

$$\mathcal{G}_{0,C,A} = C^* C. \quad (4.6)$$

We next introduce the operator map

$$\Gamma_{n,A} = (I - B_A)^n: \quad X \mapsto \sum_{k=0}^n (-1)^k \binom{n}{k} A^{*k} X A^k. \quad (4.7)$$

There is then an assortment of identities as listed below.

**Lemma 4.1.** (1) For all  $H, A \in \mathcal{L}(\mathcal{X})$  and any integers  $k \geq 1$  and  $N \geq 0$ ,

$$\Gamma_{k,A}[H] = \Gamma_{k-1,A}[H] - A^* \Gamma_{k-1,A}[H] A, \quad (4.8)$$

$$H = \sum_{j=0}^N \binom{k+j-1}{j} A^{*j} \Gamma_{k,A}[H] A^j + \sum_{j=1}^k \binom{N+k}{N+j} A^{*N+j} \Gamma_{k-j,A}[H] A^{N+j}. \quad (4.9)$$

(2) If  $(C, A)$  is  $n$ -output stable, then also

$$\mathcal{G}_{m,C,A} - A^* \mathcal{G}_{m,C,A} A = \mathcal{G}_{m-1,C,A} \quad \text{for } m = 1, \dots, n, \quad (4.10)$$

$$\Gamma_{k,A}[\mathcal{G}_{n,C,A}] = \mathcal{G}_{n-k,C,A} \quad \text{for } k = 0, \dots, n. \quad (4.11)$$

Moreover in this case  $(C, A)$  is also  $k$ -output stable for  $k = 1, \dots, n-1$ .

*Proof.* Upon applying identity  $(I - B_A)^k = (I - B_A)^{k-1} - B_A(I - B_A)^{k-1}$  to an operator  $H \in \mathcal{L}(\mathcal{X})$  and making use of definition (4.7) we get (4.8). To verify (4.9), we use the identity (2.8) in the form

$$\sum_{j=0}^N \binom{k+j-1}{j} z^j = \frac{1}{(1-z)^k} - \sum_{j=1}^k \binom{N+k}{N+j} \frac{z^{N+j}}{(1-z)^j}.$$

Multiplying both parts in the latter identity by  $(1-z)^k$  we get

$$1 = \sum_{j=0}^N \binom{k+j-1}{j} z^j (1-z)^k + \sum_{j=1}^k \binom{N+k}{N+j} z^{N+j} (1-z)^{N-j}, \quad (4.12)$$

which in turn implies the operator identity

$$I_{\mathcal{X}} = \sum_{j=0}^N \binom{k+j-1}{j} B_A^j (1 - B_A)^k + \sum_{j=1}^k \binom{N+k}{N+j} B_A^{N+j} (1 - B_A)^{N-j}.$$

Upon applying this latter identity to an operator  $H \in \mathcal{L}(\mathcal{X})$  and making use of definition (4.7) we get (4.9).

We provide two proofs of (4.10) and (4.11), one of which is a straightforward easy-to-remember computation but which is rigorous only if the spectral radius  $\rho(B_A)$  of  $B_A$  is less than 1, together with a second more elaborate proof to handle the general case. We have from (4.3) and (4.4)

$$\begin{aligned} \mathcal{G}_{n,C,A} - A^* \mathcal{G}_{m,C,A} A &= (I - B_A)[\mathcal{G}_{m,C,A}] \\ &= (I - B_A) \circ (I - B_A)^{-m} [C^* C] \\ &= (I - B_A)^{-m+1} [C^* C] = \mathcal{G}_{m-1,C,A} \end{aligned}$$

and (4.10) follows, at least for the case where  $\|B_A\| < 1$ . To handle the general case (where we only assume that  $(C, A)$  is  $n$ -output stable), one can plug in the infinite series representation (4.2) for  $\mathcal{G}_{m,C,A}$  and make use of the binomial coefficient identity  $\binom{m}{k} = \binom{m-1}{k} + \binom{m-1}{k-1}$  to arrive at the result. A corollary of the computation is that the infinite series defining  $\mathcal{G}_{m-1,C,A}$  is strongly convergent, i.e.,  $(C, A)$  is also  $(m-1)$ -output stable whenever it is  $m$ -output stable.

To prove (4.11) under the assumption that  $\rho(B_A) < 1$ , note that

$$\begin{aligned}\Gamma_{j,A}[\mathcal{G}_{n,C,A}] &= (I - B_A)^j[(I - B_A)^{-n}[C^*C]] \\ &= (I - B_A)^{-n+j}[C^*C] = \mathcal{G}_{n-j,C,A}.\end{aligned}$$

To handle the general case, one can do an inductive argument using identity (4.10) to arrive at the result. As a corollary we arrive at the final statement in Lemma 4.1:  $(C, A)$  is  $k$ -output stable for  $k = 1, \dots, n-1$  whenever  $(C, A)$  is  $n$ -output stable.  $\square$

**Definition 4.2.** The operator  $A \in \mathcal{L}(\mathcal{X})$  is called  $n$ -contractive if  $\Gamma_{n,A}[I]$  is positive semidefinite and it is called  $n$ -hypercontractive if  $\Gamma_{k,A}[I] \geq 0$  for all  $0 \leq k \leq n$ .

It was shown in [1] (see also [25] as well as [26] for a multivariable version) that inequalities  $\Gamma_{1,A}[I] \geq 0$  and  $\Gamma_{n,A}[I] \geq 0$  imply that  $A$  is an  $n$ -hypercontraction. This result extends from  $I$  to an arbitrary  $H \geq 0$ ; the proof below is modelled from the one in [25].

**Lemma 4.3.** Let us assume that the operators  $H, A \in \mathcal{L}(\mathcal{X})$  are such that

$$H \geq A^*HA \geq 0 \quad \text{and} \quad \Gamma_{n,A}[H] \geq 0 \tag{4.13}$$

for some integer  $n \geq 3$ . Then

$$\Gamma_{k,A}[H] \geq 0 \quad \text{for all } k = 1, \dots, n-1. \tag{4.14}$$

*Proof.* Observe that the leftmost inequalities in (4.13) mean that  $H$  and  $\Gamma_{1,A}[H]$  are both positive semidefinite. Making use of definitions (4.3) and (4.7) we have

$$\begin{aligned}\sum_{j=0}^N A^{*j} \Gamma_{n-1,A}[H] A^j &= \sum_{j=0}^N B_A^j (I - B_A)^{n-1}[H] \\ &= (I - B_A^{N+1})(I - B_A)^{n-2}[H] \\ &= \Gamma_{n-2,A}[H] - A^{*N+1} \Gamma_{n-2,A}[H] A^{N+1}.\end{aligned} \tag{4.15}$$

Iterating the first inequality in (4.13) gives  $A^{*j} HA^j \leq H$  for all  $j \geq 0$  and therefore,

$$|\langle \Gamma_{n-2,A}[H] A^k x, A^k x \rangle| \leq \sum_{j=0}^{n-2} \binom{n-2}{j} \langle Hx, x \rangle = 2^{n-2} \langle Hx, x \rangle. \tag{4.16}$$

Taking the inner product of both parts in (4.15) against  $x \in \mathcal{X}$  and then making use of (4.16) gives

$$\begin{aligned}\left| \sum_{j=0}^N \langle \Gamma_{n-1,A}[H] A^j x, A^j x \rangle \right| &= |\langle \Gamma_{n-2,A}[H] x, x \rangle - \langle \Gamma_{n-2,A}[H] A^{N+1} x, A^{N+1} x \rangle| \\ &\leq 2^{n-1} \langle Hx, x \rangle.\end{aligned} \tag{4.17}$$

On account of relation (4.8), the second inequality in (4.13) implies

$$\Gamma_{n-1,A}[H] \geq A^* \Gamma_{n-1,A}[H] A.$$

Therefore  $\langle \Gamma_{n-1,A}[H] A^j x, A^j x \rangle \geq \langle \Gamma_{n-1,A}[H] A^{j+1} x, A^{j+1} x \rangle$  for all  $j \geq 0$  and all  $x \in \mathcal{X}$ . Thus, on the left hand side of (4.17) we have the partial sum of a non-increasing sequence and, since the partial sums are uniformly bounded (by  $2^{n-1} \langle Hx, x \rangle$ ), it follows that all the terms in the sequence are nonnegative. In particular,  $\langle \Gamma_{n-1,A}[H] x, x \rangle \geq 0$ . Since the latter inequality holds for every  $x \in \mathcal{X}$ , we conclude that  $\Gamma_{n-1,A}[H] \geq 0$ . We then obtain recursively all the desired inequalities in (4.14).  $\square$

**Definition 4.4.** The operator  $A \in \mathcal{L}(\mathcal{X})$  is called *strongly stable* if  $A^k$  tends to zero as  $k \rightarrow \infty$  in the strong operator topology, i.e.,  $\|A^k x\| \rightarrow 0$  for every  $x \in \mathcal{X}$ .

The following result gives connections between  $n$ -output stability, observability gramians and solutions of associated Stein equations and inequalities. In case  $n = 1$  it is well-known. In what follows, we will refer to the last relations in (4.18) and (4.19) as the *Stein inequality* and the *Stein equality*, respectively.

**Theorem 4.5.** Let  $C \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$  and  $A \in \mathcal{L}(\mathcal{X})$ . Then:

- (1) The pair  $(C, A)$  is  $n$ -output-stable if and only if there exists an  $H \in \mathcal{L}(\mathcal{X})$  satisfying the system of inequalities

$$H \geq A^* H A \geq 0 \quad \text{and} \quad \Gamma_{n,A}[H] \geq C^* C. \quad (4.18)$$

- (2) If  $(C, A)$  is  $n$ -output-stable, then the observability gramian  $\mathcal{G}_{n,C,A}$  satisfies the system

$$H \geq A^* H A \geq 0 \quad \text{and} \quad \Gamma_{n,A}[H] = C^* C \quad (4.19)$$

and is the minimal positive semidefinite solution of the system (4.18).

- (3) There is a unique positive-semidefinite solution  $H$  of the system (4.19) with  $H = \mathcal{G}_{n,C,A}$  if  $A$  is strongly stable. If  $A$  is a contraction, then the solution of the system (4.19) is unique if and only if  $A$  is strongly stable.

*Proof.* Suppose first that  $(C, A)$  is  $n$ -output-stable. Then the infinite series in (4.2) converges in the strong operator topology to the operator  $H = \mathcal{G}_{n,C,A} \geq 0$  and by (4.10), (4.11) and (4.6),

$$\mathcal{G}_{n,C,A} - A^* \mathcal{G}_{n,C,A} A = \mathcal{G}_{n-1,C,A} \geq 0, \quad \Gamma_{n,A}[\mathcal{G}_{n,C,A}] = \mathcal{G}_{0,C,A} = C^* C.$$

Thus,  $H = \mathcal{G}_{n,C,A}$  satisfies relations (4.19) and hence also inequalities (4.18).

Conversely, suppose that inequalities (4.18) are satisfied for some  $H \in \mathcal{L}(\mathcal{X})$ . Then for every integer  $N \geq 0$  we have

$$\begin{aligned} \sum_{j=0}^N \binom{j+n-1}{j} A^{*j} C^* C A^j &\leq \sum_{j=0}^N \binom{j+n-1}{j} A^{*j} \Gamma_{n,A}[H] A^j \\ &= H - \sum_{j=1}^n \binom{N+n}{N+j} A^{*N+j} \Gamma_{n-j,A}[H] A^{N+j} \leq H. \end{aligned} \quad (4.20)$$

Indeed, the first inequality follows since  $C^* C \leq \Gamma_{n,A}[H]$ , the second equality holds by (4.9) (with  $k = n$ ) and the last inequality holds since  $\Gamma_{k,A}[H] \geq 0$  for  $k = 0, \dots, n-1$ , by the assumptions (4.18) and Lemma 4.3.

By letting  $N \rightarrow \infty$  in (4.20) we conclude that the left-hand side sum converges (weakly and therefore, since all the terms are positive semidefinite, strongly) to a bounded positive semidefinite operator. By (4.2),

$$\lim_{N \rightarrow \infty} \sum_{j=0}^N \binom{j+n-1}{j} A^{*j} C^* C A^j = \sum_{j=0}^{\infty} \binom{j+n-1}{j} A^{*j} C^* C A^j = \mathcal{G}_{n,C,A}$$

and passing to the limit in (4.20) as  $N \rightarrow \infty$  gives  $\mathcal{G}_{n,C,A} \leq H$ . In particular the operator  $\mathcal{G}_{n,C,A}$  is bounded (since  $H$  is) and therefore the pair  $(C, A)$  is  $n$ -output-stable. This completes the proof of the first part of the theorem.

As observed above, the gramian  $\mathcal{G}_{n,C,A}$  satisfies relations (4.19) (and therefore, relations (4.18)). We also showed that any operator  $H$  solving the inequalities (4.18) also satisfies the inequality  $H \geq \mathcal{G}_{n,C,A}$  so that  $\mathcal{G}_{n,C,A}$  is indeed the minimal solution to the system of inequalities (4.18). This completes the proof of the second part of the theorem.

Now suppose that  $A$  is strongly stable and that  $H \in \mathcal{L}(\mathcal{X})$  solves the system (4.19). We will show that then necessarily  $H = \mathcal{G}_{n,C,A}$ . We first recall that if positive semidefinite operators  $P, Q \in \mathcal{L}(\mathcal{X})$  satisfy the Stein equation

$$P - A^* P A = Q \quad (4.21)$$

with a strongly stable  $A \in \mathcal{L}(\mathcal{X})$ , then  $P$  is uniquely recovered from (4.21) via the strongly converging series

$$P = \sum_{k=0}^{\infty} A^{*k} Q A^k. \quad (4.22)$$

Indeed, iterating (4.21) gives

$$P = A^{*N} P A^N + \sum_{j=0}^{N-1} A^{*j} Q A^j \quad \text{for all } N \geq 0,$$

and since all the terms in the latter equality are positive semidefinite, the convergence of the series on the right side of (4.22) follows. The strong stability of  $A$  guarantees that  $A^{*N} P A^N \rightarrow 0$  as  $N \rightarrow \infty$ , which implies equality in (4.22). Now we observe that by (4.8), the Stein equation in (4.19) can be written as

$$\Gamma_{n-1,A}[H] - A^* \Gamma_{n-1,A}[H] A = C^* C$$

and the above uniqueness shows that

$$\Gamma_{n-1,A}[H] = \sum_{k=0}^{\infty} A^{*k} C^* C A^k = \mathcal{G}_{1,C,A}.$$

The latter equality can be in turn written in the form (4.21) with  $P = \Gamma_{n-2,A}[H]$  and  $Q = \mathcal{G}_{1,C,A}$ , and by the same uniqueness argument we have

$$\Gamma_{n-2,A}[H] = \sum_{k=0}^{\infty} A^{*k} \mathcal{G}_{1,C,A} A^k = \mathcal{G}_{2,C,A}.$$

Continuing this procedure, we get  $\Gamma_{n-j,A}[H] = \mathcal{G}_{j,C,A}$  for  $j = 1, \dots, n$ . For  $j = n$  we have in particular,  $H = \Gamma_{0,A}[H] = \mathcal{G}_{n,C,A}$ , which gives the desired uniqueness.

To complete the proof of part (3) of the theorem it remains to show that if  $A$  is a contraction and the system (4.19) admits a unique solution (which necessarily is

$H = \mathcal{G}_{C,A,n}$ ), then the operator  $A$  is strongly stable. We prove the contrapositive: if  $A$  is not strongly stable, then the solution of (4.19) is not unique.

Since  $A$  is a contraction, the sequence of operators  $\Delta_k = A^{*k}A^k$  is decreasing and bounded below and therefore has a strong limit  $\Delta \geq 0$  which clearly satisfies the relation  $B_A[\Delta] = \Delta$  and which is not zero, since  $A$  is assumed not to be strongly stable. Then it follows that

$$\Gamma_{k,A}[\Delta] = (I - B_A)^k[\Delta] = \sum_{j=0}^k (-1)^j \binom{k}{j} \Delta = 0$$

for all  $k \in \{1, \dots, n\}$ . Now we see that the operator  $H = \mathcal{G}_{n,C,A} + \Delta$  (as well as  $\mathcal{G}_{n,C,A}$ ) satisfies the system (4.19) which therefore has more than one positive-semidefinite solution.  $\square$

**Definition 4.6.** A pair  $(C, A)$  with  $C \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$  and  $A \in \mathcal{L}(\mathcal{X})$  is called *n-contractive* if inequalities (4.18) hold with  $H = I_{\mathcal{X}}$ , that is, if  $A$  is *n-hypercontractive* and in addition

$$\Gamma_{n,A}[I_{\mathcal{X}}] \geq C^*C, \quad \text{or equivalently} \quad C^*C + \sum_{j=1}^n \binom{n}{j} A^{*j}A^j \leq I_{\mathcal{X}}.$$

The pair  $(C, A)$  will be called *n-isometric* if relations (4.19) hold with  $H = I_{\mathcal{X}}$ , that is, if  $A$  is *n-hypercontractive* and in addition

$$\Gamma_{n,A}[I_{\mathcal{X}}] = C^*C, \quad \text{or equivalently} \quad C^*C + \sum_{j=1}^n \binom{n}{j} A^{*j}A^j = I_{\mathcal{X}}.$$

A pair  $(C, A)$  is called *observable* if the operator  $\mathcal{O}_{1,C,A}$  (equivalently,  $\mathcal{G}_{1,C,A}$ ) is injective. This property means that a state space vector  $x \in \mathcal{X}$  is uniquely recovered from the output string  $\{y_k\}_{k \geq 0}$  generated by running the system (3.2) with the initial condition  $x_0 = x$  and the zero input string. A pair  $(C, A)$  is called *exactly observable* if  $\mathcal{O}_{1,C,A}$  (equivalently,  $\mathcal{G}_{1,C,A}$ ) is bounded and bounded from below. More generally, we say that the pair  $(C, A)$  is *exactly n-observable* if the *n*-gramian  $\mathcal{G}_{n,C,A}$  is bounded and bounded below. While *n*-observability implies *k*-observability for  $1 \leq k < n$ , the corresponding statement for exact *k*-observability fails (see Proposition 5.7 below).

The following statements follow along the lines of Theorem 4.5.

**Proposition 4.7.** (1) Suppose that  $(C, A)$  is an *n-contractive* pair. Then  $(C, A)$  is *n-output-stable* with  $\mathcal{G}_{n,C,A} \leq I_{\mathcal{X}}$  and the observability gramian  $\mathcal{G}_{n,C,A}$  is the unique solution of system (4.19) if and only if  $A$  is strongly stable.

(2) Suppose that  $(C, A)$  is an *n-isometric* pair. Then  $H = I_{\mathcal{X}}$  is the unique solution of the system (4.19) if and only if  $A$  is strongly stable. In this case  $\mathcal{O}_{n,C,A}$  is isometric and hence also  $(C, A)$  is exactly observable.

Let us say that the pair  $(C, A)$  is similar to the pair  $(\tilde{C}, \tilde{A})$  if there is an invertible operator  $T$  on  $\mathcal{X}$  so that  $\tilde{C} = CT^{-1}$  and  $\tilde{A} = TAT^{-1}$ . Then we have the following characterization of pairs  $(C, A)$  which are similar to an *n-contractive* or to an *n-isometric* pair.

**Proposition 4.8.** (1) *The pair  $(C, A)$  is similar to an  $n$ -contractive pair  $(\tilde{C}, \tilde{A})$  if and only if there exists a bounded, strictly positive-definite solution  $H$  to the system of inequalities (4.18).*

(2) *The pair  $(C, A)$  is similar to an  $n$ -isometric pair if and only if there exists a bounded, strictly positive-definite solution  $H$  of the system (4.19).*

*Proof.* Suppose that  $H$  is a strictly positive-definite solution of (4.18). Factor  $H$  as  $H = T^*T$  with  $T$  invertible and set

$$\tilde{C} = CT^{-1}, \quad \tilde{A} = TAT^{-1}. \quad (4.23)$$

Multiplying each inequality in (4.18) on the left by  $T^{*-1}$  and on the right by  $T^{-1}$  and replacing  $A$  and  $C$  respectively by  $T^{-1}\tilde{A}T$  and  $\tilde{C}T$  then leads us to

$$I_{\mathcal{X}} \geq \tilde{A}^*\tilde{A} \quad \text{and} \quad \Gamma_{n,\tilde{A}}[I_{\mathcal{X}}] \geq \tilde{C}^*\tilde{C}$$

which means that  $(\tilde{C}, \tilde{A})$  is an  $n$ -contractive pair which is similar to the original pair  $(C, A)$ . Conversely, if  $(\tilde{C}, \tilde{A})$  given by (4.23) is contractive, then  $H = T^*T$  is bounded and strictly positive-definite and satisfies the inequalities (4.18). This verifies the first statement of the Proposition. The second statement follows in a similar way.  $\square$

As a consequence of the observations in Proposition 4.8, Proposition 4.7 can be formulated more generally as follows.

**Proposition 4.9.** (1) *If the pair  $(C, A)$  is such that inequalities (4.18) have a strictly positive-definite solution  $H$ , then  $(C, A)$  is  $n$ -output-stable. Moreover, the observability gramian  $\mathcal{G}_{n,C,A}$  is the unique positive semidefinite solution of the system (4.19) if and only if  $A$  is strongly stable.*

(2) *If the pair  $(C, A)$  is such that the system (4.19) has a strictly positive-definite solution  $H$ , then  $(C, A)$  is  $n$ -output-stable and the observability gramian  $\mathcal{G}_{n,C,A}$  is the unique positive semidefinite solution of the system (4.19) if and only if  $A$  is strongly stable. In this case  $(C, A)$  is moreover exactly observable.*

The last part of Proposition 4.9 has a converse.

**Proposition 4.10.** *Suppose that the pair  $(C, A)$  is  $n$ -output-stable and exactly observable. Then  $A$  is strongly stable.*

*Proof.* Observe that if  $H$  is any solution to the system (4.19), then (4.20) takes the form

$$\sum_{j=0}^N \binom{j+n-1}{j} A^{*j} C^* C A^j = H - \sum_{j=1}^n \binom{N+n}{N+j} A^{*N+j} \Gamma_{n-j,A}[H] A^{N+j} \quad (4.24)$$

and still holds for all  $N \geq 0$ . If  $(C, A)$  is  $n$ -output-stable and exactly observable, then the observability gramian  $\mathcal{G}_{n,C,A}$  is a strictly positive-definite solution of the system (4.19). Plugging  $H = \mathcal{G}_{n,C,A}$  into (4.24) gives

$$\begin{aligned} \mathcal{G}_{n,C,A} &= \sum_{k=0}^N \binom{k+n-1}{k} A^{*k} C^* C A^k + \sum_{j=1}^n \binom{N+n}{N+j} A^{*N+j} \Gamma_{n-j,A}[\mathcal{G}_{n,C,A}] A^{N+j} \\ &= \sum_{k=0}^N \binom{k+n-1}{k} A^{*k} C^* C A^k + \sum_{j=1}^n \binom{N+n}{N+j} A^{*N+j} \mathcal{G}_{j,C,A} A^{N+j} \end{aligned} \quad (4.25)$$

where the last equality follows from (4.11). From the infinite-series representation (4.2) for  $\mathcal{G}_{C,A,n}$ , taking limits in (4.25) gives

$$\lim_{N \rightarrow \infty} \binom{N+n}{N+j} A^{*N+j} \mathcal{G}_{j,C,A} A^{N+j} = 0$$

for  $j = 1, \dots, n$ . In particular, we have

$$\lim_{N \rightarrow \infty} A^{*N+n} \mathcal{G}_{n,C,A} A^{N+n} = 0. \quad (4.26)$$

Since  $\mathcal{G}_{n,C,A}$  is strictly positive definite, we conclude that there is an  $\varepsilon > 0$  so that

$$\varepsilon \|x\|^2 \leq \langle \mathcal{G}_{n,C,A} x, x \rangle \quad \text{for all } x \in \mathcal{X}.$$

Upon combining (4.26) with the latter relations (with  $x$  replaced by  $A^N x$ ) we get

$$\varepsilon \|A^N x\|^2 \leq \langle \mathcal{G}_{n,C,A} A^N x, A^N x \rangle \rightarrow 0$$

for all  $x \in \mathcal{X}$ , and we conclude that  $A$  is strongly stable as asserted.  $\square$

We conclude this section with formal introduction of backward-shifted versions of the observability operator  $\mathcal{O}_{n,C,A}$  (3.7) and observability gramian  $\mathcal{G}_{n,C,A}$  (4.1). The latter operators can be expressed as

$$\mathcal{O}_{n,C,A} : x \rightarrow CR_n(zA)x \quad \text{and} \quad \mathcal{G}_{n,C,A} = R_n(B_A)[C^*C]$$

where  $R_n(z) = (1-z)^{-n}$  and where  $B_A$  is the operator on  $\mathcal{L}(\mathcal{X})$  given by (4.3). We now introduce the backward-shifted variants of these objects by simply replacing the function  $R_n$  in the two formulas above by its backward shifts  $R_{n,k}$ :

$$\mathfrak{O}_{n,k,C,A} : x \rightarrow CR_{n,k}(zA)x = \sum_{j=0}^{\infty} \binom{n+j+k-1}{j+k} (CA^j x) z^j, \quad (4.27)$$

$$\mathfrak{G}_{n,k,C,A} := R_{n,k}(B_A)[C^*C] = \sum_{j=0}^{\infty} \binom{n+j+k-1}{j+k} A^{j*} C^* CA^j. \quad (4.28)$$

Observe that the latter power series representations follow from that in (2.12). Letting  $k = 0$  in (4.27), (4.28) we conclude

$$\mathfrak{O}_{n,0,C,A} = \mathcal{O}_{n,C,A} \quad \text{and} \quad \mathfrak{G}_{n,0,C,A} = \mathcal{G}_{n,C,A}. \quad (4.29)$$

If  $k \geq 1$ , then we can use the formula (2.5) along with (4.27), (4.28) to get representations

$$\mathfrak{O}_{n,k,C,A} = \sum_{\ell=1}^n \binom{\ell+k-2}{\ell-1} \mathcal{O}_{n-\ell+1,C,A}, \quad \mathfrak{G}_{n,k,C,A} = \sum_{\ell=1}^n \binom{\ell+k-2}{\ell-1} \mathcal{G}_{n-\ell+1,C,A}$$

which imply in particular that for an  $n$ -output stable pair  $(C, A)$ , the operator  $\mathfrak{O}_{n,k,C,A} : \mathcal{X} \rightarrow \mathcal{A}_n(\mathcal{Y})$  is bounded for all  $k \geq 0$ .

**Proposition 4.11.** *The weighted Stein identity*

$$A^* \mathfrak{G}_{n,k+1,C,A} A + \binom{n+k-1}{k} \cdot C^* C = \mathfrak{G}_{n,k,C,A} \quad (4.30)$$

holds for all integers  $k \geq 0$ .

*Proof.* The operatorial equality

$$R_{n,k}(B_A) - B_A \circ R_{n,k+1}(B_A) = R_{n,k}(0) \cdot I_{\mathcal{L}(\mathcal{X})} = \binom{n+k-1}{k} \cdot I_{\mathcal{L}(\mathcal{X})}$$

is a consequence of the identity (2.7). Applying this equality to the operator  $C^*C$  gives, on account of (4.28),

$$\begin{aligned} (R_{n,k}(B_A) - B_A \circ R_{n,k+1}(B_A)) [C^*C] &= \mathfrak{G}_{n,k,C,A} - A^* \mathfrak{G}_{n,k+1,C,A} A \\ &= \binom{n+k-1}{k} \cdot C^*C \end{aligned}$$

and we arrive at (4.30) as wanted.  $\square$

As a consequence of Proposition 2.3, it turns out that  $\mathfrak{G}_{n,k,C,A}$  can be expressed in terms of  $\mathcal{G}_{n,C,A}$  as follows.

**Proposition 4.12.** *The operator  $\mathfrak{G}_{n,k,C,A}$  given by (4.28) admits the representation*

$$\mathfrak{G}_{n,k,C,A} = \sum_{\kappa=0}^{n-1} \left( \sum_{j=0}^{\kappa} \binom{k+n-1}{k} \binom{n-1-\kappa+j}{j} (-1)^j \right) A^{*\kappa} \mathcal{G}_{n,C,A} A^{\kappa}. \quad (4.31)$$

*Proof.* Note that the identity (2.9) leads to the operator identity

$$R_{n,k}(B_A) = \sum_{\kappa=0}^{n-1} \left( \sum_{j=0}^{\kappa} \binom{k+n-1}{k} \binom{n-1-\kappa+j}{j} (-1)^j \right) (B_A)^{\kappa} \circ R_n(B_A).$$

Application of this operator to  $C^*C$  now leads to the identity (4.31) due to (4.28), (4.3) and (4.4).  $\square$

## 5. OBSERVABILITY-OPERATOR RANGE SPACES AND REPRODUCING KERNEL HILBERT SPACES

Let  $S_n$  denote the shift operator on the space  $\mathcal{A}_n(\mathcal{Y})$  defined as  $S_n : f(z) \rightarrow zf(z)$ . Reproducing kernel calculations show that its adjoint  $S_n^*$  is given by

$$S_n^* f = \sum_{j=0}^{\infty} \frac{j+1}{n+j} \cdot f_{j+1} z^j = \sum_{j=0}^{\infty} \frac{\mu_{n,j+1}}{\mu_{n,j}} \cdot f_{j+1} z^j \quad \text{if } f(z) = \sum_{j=0}^{\infty} f_j z^j. \quad (5.1)$$

Iterating the latter formula gives

$$S_n^{*k} f = \sum_{j=0}^{\infty} \frac{(k+j)!(n+j-1)!}{j!(n+k+j-1)!} \cdot f_{k+j} z^j = \sum_{j=0}^{\infty} \frac{\mu_{n,j+k}}{\mu_{n,j}} \cdot f_{k+j} z^j. \quad (5.2)$$

The operator  $S_n^*$  is a strongly stable  $n$ -hypercontraction. As was shown in [28, Lemma 5.1] for  $f$  of the form (5.1),

$$\sum_{j=0}^n (-1)^j \binom{n}{j} \|S_n^{*j} f\|_{\mathcal{A}_n}^2 = \|f_0\|^2. \quad (5.3)$$

The latter equality means that the pair  $(E, S_n^*)$  is  $n$ -isometric in the sense of Definition 4.6, where  $E : \mathcal{A}_n(\mathcal{Y}) \rightarrow \mathcal{Y}$  is the evaluation operator defined by  $Ef = f(0)$ . It follows from (5.2) that

$$ES_n^{*k} f = \frac{j!(n-1)!}{(n+k-1)!} f_k = \mu_{n,k} f_k \quad (5.4)$$

and therefore,

$$\mathcal{O}_{n,E,S_n^*}f = E(I - zS_n^*)^{-n}f = \sum_{j=0}^{\infty} \binom{j+n-1}{j} (ES_n^{*j}f) z^j = \sum_{j=0}^{\infty} f_j z^j = f$$

so that the observability operator  $\mathcal{O}_{n,E,S_n^*}$  equals the identity operator on  $\mathcal{A}_n(\mathcal{Y})$ .

Associated with an  $n$ -output-stable pair  $(C, A)$  is the range of the observability operator

$$\text{Ran } \mathcal{O}_{n,C,A} = \{C(I - zA)^{-n}x : x \in \mathcal{X}\}.$$

**Theorem 5.1.** *Suppose that  $(C, A)$  is an  $n$ -output-stable pair. Then:*

- (1) *The intertwining relation*

$$S_n^* \mathcal{O}_{n,C,A} = \mathcal{O}_{n,C,A} A \quad (5.5)$$

*holds and hence the linear manifold  $\mathcal{M} = \text{Ran } \mathcal{O}_{n,C,A}$  is  $S_n^*$ -invariant.*

- (2) *Let  $H \in \mathcal{L}(\mathcal{X})$  satisfy inequalities (4.18) and let  $\mathcal{X}'$  be the completion of  $\mathcal{X}$  with inner product  $\|[x]\|_{\mathcal{X}'}^2 = \langle Hx, x \rangle_{\mathcal{X}}$  (where  $[x]$  denotes the equivalence class modulo  $\text{Ker } H$  generated by  $x$ ). Then  $A$  and  $C$  extend to define bounded operators  $A' : \mathcal{X}' \rightarrow \mathcal{X}'$  and  $C' : \mathcal{X}' \rightarrow \mathcal{Y}$  and the observability operator  $\mathcal{O}_{n,C,A}$  extends to define a contraction operator  $\mathcal{O}_{n,C',A'} : \mathcal{X}' \rightarrow \mathcal{A}_n(\mathcal{Y})$ . Moreover,  $\mathcal{O}_{n,C',A'}$  is an isometry if and only if  $H$  satisfies relations (4.19) and  $A'$  is strongly stable, i.e.,  $\langle HA^N x, A^N x \rangle \rightarrow 0$  for all  $x \in \mathcal{X}$  as  $N \rightarrow \infty$ .*

- (3) *If the linear manifold  $\mathcal{M} := \text{Ran } \mathcal{O}_{n,C,A}$  is given the lifted norm*

$$\|\mathcal{O}_{n,C,A}x\|_{\mathcal{M}}^2 = \inf_{y \in \mathcal{X} : \mathcal{O}_{n,C,A}y = \mathcal{O}_{n,C,A}x} \langle Hy, y \rangle_{\mathcal{X}},$$

*then*

- (a)  *$\mathcal{M}$  can be completed to  $\mathcal{M}' = \text{Ran } \mathcal{O}_{n,C',A'}$  with contractive inclusion in  $\mathcal{A}_n(\mathcal{Y})$ :*

$$\|f\|_{\mathcal{A}_n}^2 \leq \|f\|_{\mathcal{M}'}^2 \quad \text{for all } f \in \mathcal{M}'.$$

*Furthermore,  $\mathcal{M}'$  is isometrically equal to the reproducing kernel Hilbert space with reproducing kernel  $K_{n,C,A,H}$  given by*

$$K_{n,C,A,H}(z, \zeta) = C(I - zA)^{-n}H(I - \bar{\zeta}A^*)^{-n}C^*. \quad (5.6)$$

- (b) *The operator  $S_n^* : \mathcal{M} \rightarrow \mathcal{M}$  defined as in (5.1) is a contraction and*

$$\sum_{j=0}^n (-1)^j \binom{n}{j} \|S_n^{*j}f\|_{\mathcal{M}}^2 \geq \|f(0)\|_{\mathcal{Y}}^2 \quad \text{for all } f \in \mathcal{M}. \quad (5.7)$$

*Moreover, if the Stein equality in (4.19) holds, then (5.7) holds with equality.*

- (4) *Conversely, if  $\mathcal{M}$  is a Hilbert space contractively included in  $\mathcal{A}_n(\mathcal{Y})$  which is invariant under  $S_n^*$  which in turn is a contraction on  $\mathcal{M}$  and for which (5.7) holds, then there is an  $n$ -contractive pair  $(C, A)$  such that  $\mathcal{M} = \mathcal{H}(K_{n,C,A,I}) = \text{Ran } \mathcal{O}_{n,C,A}$  isometrically. In case (5.7) holds with equality, then  $(C, A)$  can be taken to be  $n$ -isometric. The canonical-model choice of such a pair  $(C, A)$  is*

$$(C, A) = (E|_{\mathcal{M}}, S_n^*|_{\mathcal{M}}) \text{ where } E : f \in \mathcal{A}_n(\mathcal{Y}) \mapsto f(0).$$

*Proof of (1):* Making use of power series expansion (2.3) and of (5.1) we get (5.5):

$$\begin{aligned} S_n^* \mathcal{O}_{C,A,n} x &= S_n^*(C(I-zA)^{-n}x) = \sum_{k=0}^{\infty} \frac{k+1}{n+k} \cdot \binom{n+k}{k+1} (CA^{k+1}x)z^k \\ &= \sum_{k=0}^{\infty} \binom{n+k-1}{k} (CA^{k+1}x)z^k \\ &= C(I-zA)^{-n}Ax = \mathcal{O}_{C,A,n}Ax. \end{aligned}$$

*Proof of (2):* Inequalities (4.18) amount to the statement that the pair  $(C, A)$  is  $n$ -contractive and well-defined on the dense subset  $[\mathcal{X}]$  of  $\mathcal{X}'$  (where  $[x]$  is the equivalence class containing  $x$ ) and hence extends to an  $n$ -contractive pair  $(C', A')$  on all of  $\mathcal{X}'$  and moreover relations (4.20) hold for all  $N \geq 0$ . By Theorem 4.5,  $\mathcal{G}_{n,C,A}x \leq H$  and therefore

$$\|\mathcal{O}_{n,C,A}x\|_{\mathcal{A}_n(\mathcal{Y})} = \langle \mathcal{G}_{n,C,A}x, x \rangle_{\mathcal{X}} \leq \langle Hx, x \rangle_{\mathcal{X}},$$

so  $\mathcal{O}_{n,C,A}$  is contractive from  $\mathcal{X}$  to  $\mathcal{A}_n(\mathcal{Y})$ . It follows from (4.20) that the limit

$$\Delta_{H,A} = \lim_{N \rightarrow \infty} \sum_{j=1}^n \binom{N+n}{N+j} A^{*N+j} \Gamma_{n-j,A}[H] A^{N+j} \quad (5.8)$$

exists in the strong sense. We have from (4.20)

$$\begin{aligned} 0 &\leq \sum_{j=0}^N \binom{j+n-1}{j} A^{*j} (\Gamma_{n,A}[H] - C^*C) A^j \\ &\leq H - \sum_{j=1}^n \binom{N+n}{N+j} A^{*N+j} \Gamma_{n-j,A}[H] A^{N+j} - \sum_{j=0}^N \binom{j+n-1}{j} A^{*j} C^*C A^j \end{aligned}$$

and letting  $N \rightarrow \infty$  we get, on account of (5.8) and (4.2),

$$0 \leq \sum_{j=0}^{\infty} \binom{j+n-1}{j} A^{*j} (\Gamma_{n,A}[H] - C^*C) A^j \leq H - \Delta_{H,A} - \mathcal{G}_{n,C,A}. \quad (5.9)$$

By definition,  $\mathcal{O}_{n,C',A'}: \mathcal{X}' \rightarrow \mathcal{A}_n(\mathcal{Y})$  being an isometry means that  $\mathcal{G}_{n,C,A} = H$  in which case (5.9) forces  $\Delta_{H,A} = 0$  and equalities throughout (5.9). Since all the terms on the right hand side of (5.8) are positive semidefinite, the condition  $\Delta_{H,A} = 0$  implies in particular that

$$0 = \lim_{N \rightarrow \infty} A^{*N+n} \Gamma_{0,A}[H] A^{N+n} = \lim_{N \rightarrow \infty} A^{*N+n} H A^{N+n} = \lim_{N \rightarrow \infty} A^{*N} H A^N,$$

so that  $A'$  is strongly stable. All the terms in the series in (5.9) are nonnegative and therefore, each term equals zero. The term corresponding to the index  $j = 0$  is  $\Gamma_{n,A}[H] - C^*C$ . Hence  $\Gamma_{n,A}[H] = C^*C$  and thus the Stein equality in (4.19) holds. Conversely, if  $A'$  is strongly stable and relations (4.19) hold, we conclude as in the proof of Theorem 4.5 that  $\mathcal{G}_{n,C',A'} = H$ , i.e., that  $\mathcal{O}_{n,C',A'}$  is an isometry from  $\mathcal{X}'$  into  $\mathcal{A}_n(\mathcal{Y})$ .

*Proof of (3a):* Statement (3a) follows from general principles laid out in [9] (see also [10, 27] for applications very close to the context here). For the sake of completeness

and since the same construction arises again in the sequel, we sketch the argument here. Note that for  $x \in \mathcal{X}$  and  $y \in \mathcal{Y}$  we have

$$\begin{aligned} \langle (\mathcal{O}_{n,C,A}x)(\zeta), y \rangle_{\mathcal{Y}} &= \langle C(I - \zeta A)^{-n}x, y \rangle_{\mathcal{Y}} \\ &= \langle x, H(I - \bar{\zeta}A^*)^{-n}C^*y \rangle_{\mathcal{X}'} \\ &= \langle C(I - \cdot A)^{-n}x, C(I - \cdot A)^{-n}H(I - \bar{\zeta}A^*)^{-n}C^*y \rangle_{\mathcal{M}} \end{aligned}$$

where the last step follows from the assumption that  $\mathcal{O}_{n,C,A}$  is isometric from  $\mathcal{X}'$  into  $\mathcal{M}$ . From this computation we see that

$$\langle (\mathcal{O}_{n,C,A}x)(\zeta), y \rangle_{\mathcal{Y}} = \langle \mathcal{O}_{n,C,A}x, K_{n,C,A,H}(\cdot, \zeta)y \rangle_{\mathcal{M}}$$

and it follows that  $K_{n,C,A,H}$  is the reproducing kernel for  $\mathcal{M}$  as asserted.

*Proof of (3b):* For  $f$  of the form  $f(z) = C(I - zA)^{-n}x$ , we have

$$\|f\|_{\mathcal{H}(K_{n,C,A,H})}^2 = \langle Hx, x \rangle_{\mathcal{X}} \quad \text{and} \quad f(0) = Cx.$$

Now it follows from (5.5) that

$$\|S_n^{*j}f\|_{\mathcal{H}(K_{n,C,A,H})}^2 = \langle HA^jx, A^jx \rangle_{\mathcal{X}} \quad \text{for } j \geq 1.$$

With these substitutions, we see that  $S_n^*$  is a contraction on  $\mathcal{M}$  if and only if  $\langle Hx, x \rangle_{\mathcal{X}} \geq \langle HAx, Ax \rangle_{\mathcal{X}}$  for all  $x \in \mathcal{X}$  or, in operator form,  $H \geq A^*HA$ , and that on another hand, inequality (5.7) is equivalent to

$$\sum_{j=0}^n (-1)^j \binom{k}{j} \langle HA^jx, A^jx \rangle_{\mathcal{X}} \geq \|Cx\|_{\mathcal{Y}}^2 \quad \text{for all } x \in \mathcal{X},$$

or in operator form,

$$\sum_{j=0}^n (-1)^j \binom{k}{j} A^{*j}HA^j \geq C^*C \tag{5.10}$$

with equality in (5.7) equivalent to equality in (5.10). To complete verification of part (3b), it remains to remark that relation (5.10) coincide with that in (4.18) (or with that in (4.19) in case equality holds in (5.10)).

*Proof of (4):* Suppose that  $\mathcal{M}$  is a Hilbert space included in  $\mathcal{A}_n(\mathcal{Y})$  which is invariant under  $S_n^*$  which is contractive on  $\mathcal{M}$  and let us assume that the inequality (5.7) holds. Let  $A = S_n^*|_{\mathcal{M}}$  and let  $C$  be defined by  $Cf = f(0)$  for all  $f \in \mathcal{M}$ . In other words,  $C = E|_{\mathcal{M}}$ . Inequalities  $S_nS_n^* \leq I_{\mathcal{M}}$  and (5.7) mean that the pair  $(C, A)$  defined this way is  $n$ -contractive. Moreover, if equality holds in (5.7) for every  $f \in \mathcal{M}$ , then the pair  $(C, A)$  is  $n$ -isometric. It is readily seen that  $\mathcal{O}_{n,C,A} = \mathcal{O}_{n,E,S_n^*}|_{\mathcal{M}} = I_{\mathcal{M}}$ . Therefore, for each  $f \in \mathcal{M}$  we have  $\|f\|_{\mathcal{H}(K_{n,C,A,I})} = \|f\|_{\mathcal{M}}$  and thus  $\mathcal{M} = \mathcal{H}(K_{n,C,A,I})$  isometrically. It then follows from part (3a) of the theorem that in fact  $\mathcal{M}$  is contractively included in  $\mathcal{A}_n(\mathcal{Y})$ .  $\square$

The following corollary is a simple, useful special case of the general situation laid out in Theorem 5.1.

**Corollary 5.2.** *Suppose that the  $n$ -output stable pair  $(C, A)$  is also exactly  $n$ -observable. Then there is an  $S_n^*$ -invariant subspace  $\mathcal{M} \subset \mathcal{A}_n(\mathcal{Y})$  so that  $(C, A)$  is similar to the  $n$ -isometric pair  $(\tilde{C}, \tilde{A})$  given by*

$$(\tilde{C}, \tilde{A}) = (E|_{\mathcal{M}}, S_n^*|_{\mathcal{M}}), \quad \text{where } E: f \in \mathcal{A}_n(\mathcal{Y}) \mapsto f(0).$$

*Proof.* One can simply check directly that  $\mathcal{O}_{n,C,A} : \mathcal{X} \rightarrow \mathcal{A}_n(\mathcal{Y})$  considered as a linear transformation from  $\mathcal{X}$  onto  $\mathcal{M} := \text{Ran } \mathcal{O}_{n,C,A}$  provide the required similarity transformation.  $\square$

As explained by part (4) of Theorem 5.1, for purposes of study of contractively-included,  $S_n^*$ -invariant subspaces of  $\mathcal{A}_n(\mathcal{Y})$  which satisfy inequality (5.7), without loss of generality we may suppose at the start that we are working with  $\mathcal{X}'$  as the original state space  $\mathcal{X}$  and with the solution  $H$  of inequalities (4.18) to be normalized to  $H = I_{\mathcal{X}}$ . Then certain simplifications occur in parts (1)-(4) of Theorem 5.1 as explained in the next result.

**Theorem 5.3.** *Let  $(C, A)$  be an  $n$ -contractive pair with  $C \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$  and  $A \in \mathcal{L}(\mathcal{X})$ . Then:*

- (1)  *$(C, A)$  is  $n$ -output-stable and the intertwining relation (5.5) holds. Hence  $\text{Ran } \mathcal{O}_{n,C,A}$  is  $S_n^*$ -invariant.*
- (2) *The operator  $\mathcal{O}_{n,C,A} : \mathcal{X} \rightarrow \mathcal{A}_n(\mathcal{Y})$  is a contraction. Moreover  $\mathcal{O}_{n,C,A}$  is isometric if and only if  $(C, A)$  is an  $n$ -isometric pair and  $A$  is strongly stable.*
- (3) *If the linear manifold  $\mathcal{M} := \text{Ran } \mathcal{O}_{n,C,A}$  is given the lifted norm*

$$\|\mathcal{O}_{n,C,A}x\|_{\mathcal{M}} = \|Qx\|_{\mathcal{X}} \quad (5.11)$$

*where  $Q$  is the orthogonal projection of  $\mathcal{X}$  onto  $(\text{Ker } \mathcal{O}_{n,C,A})^\perp$ , then  $\mathcal{O}_{n,C,A}$  is a coisometry of  $\mathcal{X}$  onto  $\mathcal{M}$ . Moreover,  $\mathcal{M}$  is contained contractively in  $\mathcal{A}_n(\mathcal{Y})$  and is isometrically equal to the reproducing kernel Hilbert space  $\mathcal{H}(K_{n,C,A})$  with reproducing kernel  $K_{n,C,A}(z, \zeta)$  given by*

$$K_{n,C,A}(z, \zeta) = C(I - zA)^{-n}(I - \bar{\zeta}A^*)^{-n}C^*.$$

- (4) *If  $\mathcal{M} = \text{Ran } \mathcal{O}_{n,C,A} = \mathcal{H}(K_{n,C,A})$  is given the lifted norm  $\|\cdot\|_{\mathcal{H}(K_{n,C,A})}$  as in (5.11), then  $S_n^*|_{\mathcal{M}}$  is a contraction and*

$$\sum_{j=0}^n (-1)^j \binom{k}{j} \cdot \|S_n^{*j}f\|_{\mathcal{H}(K_{n,C,A})}^2 \geq \|f(0)\|_{\mathcal{Y}}^2 \quad \text{for all } f \in \mathcal{H}(K_{n,C,A}). \quad (5.12)$$

*Moreover, (5.12) holds with equality if and only the orthogonal projection  $Q$  of  $\mathcal{X}$  onto  $(\text{Ker } \mathcal{O}_{n,C,A})^\perp$  is subject to relations*

$$Q \geq A^*QA \quad \text{and} \quad \Gamma_{n,A}[Q] = C^*C. \quad (5.13)$$

*In particular, if  $(C, A)$  is observable, then (5.12) holds with equality if and only if  $(C, A)$  is an  $n$ -isometric pair.*

*Proof.* Statements (1)–(3) and all but the last part of statement (4) are direct specializations to the case  $H = I_{\mathcal{X}}$  of the corresponding results in Theorem 5.1. It remains only to analyze the conditions for equality in (5.12).

From the intertwining relation (5.5), we see that inequality  $S_n S_n^* \leq I_{\mathcal{M}}$  and the equality in (5.12) for a generic element  $f = \mathcal{O}_{n,C,A}x \in \mathcal{M}$  mean that

$$\begin{aligned} \|\mathcal{O}_{n,C,A}x\|_{\mathcal{H}(K_{n,C,A})} &\geq \|\mathcal{O}_{n,C,A}Ax\|_{\mathcal{H}(K_{n,C,A})}, \\ \sum_{j=0}^n (-1)^j \binom{k}{j} \cdot \|\mathcal{O}_{n,C,A}A^jx\|_{\mathcal{H}(K_{n,C,A})}^2 &= \|Cx\|_{\mathcal{Y}}^2 \end{aligned}$$

for all  $x \in \mathcal{X}$ . By the definition (5.11) of the  $\mathcal{H}(K_{n,C,A})$ -norm, the latter relations can be written as

$$\|Qx\|_{\mathcal{X}} \geq \|QAx\|_{\mathcal{X}} \quad \text{and} \quad \sum_{j=0}^n (-1)^j \binom{k}{j} \cdot \|QA^j x\|_{\mathcal{X}}^2 = \|Cx\|_{\mathcal{Y}}^2$$

which in turn is finally equivalent to (5.13).  $\square$

**Remark 5.4 (Variations on a theme: shift-invariance of the range of an observability operator).** A fundamental observation for us is the fact that  $\text{Ran } \mathcal{O}_{n,C,A}$  is invariant under the  $S_n^*$  due to the intertwining condition (5.5). However, for the case of a backward-shifted observability operator  $\mathfrak{D}_{n,k,C,A}$ , we see that  $\text{Ran } \mathfrak{D}_{n,k,C,A}$  is not  $S_n^*$ -invariant; a partial substitute is the relation

$$S_1^* \mathfrak{D}_{n,k,C,A} = \mathfrak{D}_{n,k+1,C,A} A$$

which follows directly from (2.14). We now list some additional curious connections between the ranges of  $S_n^k \mathfrak{D}_{n,k,C,A}$  and the operators  $S_n$  and  $S_n^*$ ; we shall not have need of any these in the sequel.

**Proposition 5.5.** *If the pair  $(C, A)$  is  $n$ -output stable, then the following relations hold:*

$$S_n^* S_n^k \mathfrak{D}_{n,k,C,A} = S_n^{k-1} \mathfrak{D}_{n,k-1,C,A} \quad \text{for all } k \geq 1. \quad (5.14)$$

$$S_n^{*m} S_n^k \mathfrak{D}_{n,k,C,A} = \begin{cases} S_n^{k-m} \mathfrak{D}_{n,k-m,C,A} & \text{if } m < k, \\ \mathcal{O}_{n,C,A} A^{m-k} & \text{if } m \geq k. \end{cases} \quad (5.15)$$

$$S_n (S_n^* S_n)^{-1} S_n^{k-1} \mathfrak{D}_{n,k-1,C,A} = S_n^k \mathfrak{D}_{n,k,C,A} \quad \text{for all } k \geq 1. \quad (5.16)$$

*Proof.* To prove (5.14), we note that by (2.3),

$$(S_n^k R_{n,k})(z) = \sum_{j=0}^{\infty} \binom{n+j+k-1}{j+k} \cdot z^{j+k}$$

and therefore, we have from (5.1)

$$\begin{aligned} S_n^* S_n^k R_{n,k}(z) &= \sum_{j=0}^{\infty} \frac{k+j}{n+k+j-1} \cdot \binom{n+k+j-1}{j+k} \cdot z^{k+j-1} \\ &= \sum_{j=0}^{\infty} \binom{n+k+j-2}{j+k-1} \cdot z^{k+j-1} = z^{k-1} R_{n,k-1}(z) = S_n^{k-1} R_{n,k-1}(z) \end{aligned}$$

which together with (4.27) implies (5.14).

The identity (5.15) follows recursively from (5.14) and (5.5). Note that the special case  $k = 0, m = 1$  gives us back (5.5). To verify (5.16), we first note that

$$S_n^* S_n : \sum_{j=0}^{\infty} f_j z^j \mapsto \sum_{j=0}^{\infty} \frac{j+1}{j+n} \cdot f_j z^j$$

from which it follows that  $S_n^* S_n$  is invertible on  $\mathcal{A}_n(\mathcal{Y})$  and moreover,

$$S_n (S_n^* S_n)^{-1} : \sum_{j=0}^{\infty} f_j z^j \mapsto \sum_{j=0}^{\infty} \frac{j+n}{j+1} \cdot f_j z^{j+1}.$$

We now note that (5.16) is an immediate consequence of (5.14):

$$\begin{aligned} S_n(S_n^*S_n)^{-1}S_n^{k-1}\mathfrak{D}_{n,k-1,C,A} &= S_n(S_n^*S_n)^{-1}S_n^*S_n^k\mathfrak{D}_{n,k,C,A} \\ &= S_nS_n^{k-1}\mathfrak{D}_{n,k,C,A} = S_n^k\mathfrak{D}_{n,k,C,A}. \end{aligned}$$

□

Finally we note that the operator  $S_n(S_n^*S_n)^{-1}$  was introduced and studied in the general context of a left-invertible Hilbert space operator  $S_n$  by Shimorin [31] where it is called the *Cauchy dual* of the left invertible operator  $S_n$ . For the latest developments, see [14] and the references there.

Just as in the case of unshifted observability operators, it is possible to represent the range of a  $k$ -shifted observability operator as a reproducing kernel Hilbert space.

**Theorem 5.6.** *The reproducing kernel for the space  $S_n^k \text{Ran } \mathfrak{D}_{n,k,C,A}$  (with inner product induced by  $\mathcal{A}_n(\mathcal{Y})$ ) is given by*

$$\mathfrak{K}_k(z, \zeta) = z^k \bar{\zeta}^k C R_{n,k}(zA) \mathfrak{G}_{n,k,C,A}^{-1} R_{n,k}(\zeta A)^* C^*. \quad (5.17)$$

*Proof.* Observe that by (4.27),

$$S_n^k \mathfrak{D}_{n,k,C,A} x = \sum_{j=k}^{\infty} \binom{n+j-1}{j} (CA^{j-k}x) z^j.$$

It then follows from (1.11) and (4.28) that

$$\begin{aligned} \|S_n^k \mathfrak{D}_{n,k,C,A} x\|_{\mathcal{A}_n(\mathcal{Y})}^2 &= \left\langle \sum_{j=k}^{\infty} \binom{n+j-1}{j} (A^*)^{j-k} C^* C A^{j-k} x, x \right\rangle_{\mathcal{X}} \\ &= \left\langle \sum_{j=0}^{\infty} \binom{n+j+k-1}{j+k} (A^*)^j C^* C A^j x, x \right\rangle_{\mathcal{X}} \\ &= \langle \mathfrak{G}_{n,k,C,A} x, x \rangle_{\mathcal{X}}. \end{aligned} \quad (5.18)$$

Therefore, by the general principle from [9] (see the proof of part (3a) of Theorem 5.1 above), it follows that the reproducing kernel for  $S_n^k \text{Ran } \mathfrak{D}_{n,k,C,A}$  is given by (5.17) as asserted. □

With all this model machinery in hand, we are now able to give the additional properties concerning  $n$ -observability and exact  $n$ -observability.

**Proposition 5.7.** *Let  $(C, A)$  be an  $n$ -output stable pair.*

- (1) *If  $(C, A)$  is  $n$ -observable, then  $(C, A)$  is also  $k$ -observable for  $1 \leq k < n$ .*
- (2) *It can happen that  $(C, A)$  is exactly  $n$ -observable but there is a  $k$  with  $1 \leq k < n$  such that  $(C, A)$  is not exactly  $k$ -observable.*
- (3) *If  $(C, A)$  is exactly  $n$ -observable (so  $\mathfrak{G}_{n,0,C,A} = \mathcal{G}_{n,C,A}$  is strictly positive definite), then  $\mathfrak{G}_{n,k,C,A}$  is strictly positive definite for  $k = 1, 2, 3, \dots$ .*

*Proof of (1):* It is not difficult to check by making use of the formula (5.1) that the adjoint Bergman shift  $S_n^*$  has no nontrivial isometric vectors for  $n \geq 2$ : if  $n \geq 2$  and  $f \in \mathcal{A}_n(\mathcal{Y})$  satisfies  $\|S_n^*f\| = \|f\|$ , then  $f = 0$ . Making use of the intertwining relation (5.5) and the weighted Stein identity (4.10), one can show: for  $n \geq 2$ , if  $(C, A)$  is  $n$ -observable, then  $(C, A)$  is also  $(n-1)$ -observable. Then a simple induction argument leads to statement (1) in the proposition.

*Proof of (2):* We let  $E: \mathcal{A}_n(\mathcal{Y}) \rightarrow \mathcal{Y}$  be the evaluation-at-zero map  $E: f \mapsto f(0)$ . The pair  $(E, S_n^*)$  is  $n$ -observable since the kernel of  $\mathcal{O}_{n,E,S_n^*}$  is trivial and it is exactly  $n$ -observable since  $\mathcal{G}_{n,E,S_n^*} = I_{\mathcal{A}_n(\mathcal{Y})}$  is strictly positive definite. However, we have from (4.10)

$$\mathcal{G}_{n-1,E,S_n^*} = \mathcal{G}_{n,E,S_n^*} - S_n \mathcal{G}_{n,E,S_n^*} S_n^* = I - S_n S_n^*$$

so that, according to (5.1),  $\mathcal{G}_{n-1,E,S_n^*}: z^j \mapsto \frac{n-1}{n+j-1} z^j$ . Thus the  $(n-1)$ -gramian  $\mathcal{G}_{n-1,E,S_n^*}$  is not strictly positive definite so that the pair  $(E, S_n^*)$  is not  $(n-1)$ -exactly observable.

*Proof of (3):* If  $(C, A)$  is exactly  $n$ -observable, then Corollary 5.2 tells us that  $(C, A)$  is similar to a model pair  $(\tilde{C}, \tilde{A}) := (E|_{\mathcal{M}}, S_n^*|_{\mathcal{M}})$  for an  $S_n^*$ -invariant subspace  $\mathcal{M} \subset \mathcal{A}_n(\mathcal{Y})$ . If  $(C, A)$  and  $(\tilde{C}, \tilde{A})$  are similar output pairs, then it is easily verified that the associated gramians  $\mathfrak{G}_{n,k,C,A} = R_{n,k}(B_A)[C^* C]$  and  $\mathfrak{G}_{n,k,\tilde{C},\tilde{A}} = R_{n,k}(B_{\tilde{A}})[\tilde{C}^* \tilde{C}]$  are congruent. Thus to show that  $\mathfrak{G}_{n,k,C,A}$  is strictly positive definite, it suffices to show that  $\mathfrak{G}_{n,k,\tilde{C},\tilde{A}}$  is strictly positive definite. For the model pair  $(\tilde{C}, \tilde{A})$  one can check that

$$\mathfrak{G}_{n,k,\tilde{C},\tilde{A}} = P_{\mathcal{M}} \mathfrak{G}_{n,k,E,S_n^*}|_{\mathcal{M}}.$$

Thus strict positive-definiteness of  $\mathfrak{G}_{n,k,\tilde{C},\tilde{A}}$  follows from strict positive-definiteness of  $\mathfrak{G}_{n,k,E,S_n^*}$ . Observe that in view of (4.27) and (5.4), we have

$$\mathfrak{D}_{n,k,E,S_n^*} f = \sum_{j=0}^{\infty} \binom{j+n+k-1}{j+k} (E S_n^{*j} f) z^j = \sum_{j=0}^{\infty} \frac{\mu_{n,j}}{\mu_{n,j+k}} f_j z^j. \quad (5.19)$$

Therefore,

$$S_n^k \mathfrak{D}_{n,k,E,S_n^*} f = \sum_{j=0}^{\infty} \frac{\mu_{n,j}}{\mu_{n,j+k}} f_j z^{j+k}$$

and by computation (5.18), we have

$$\langle \mathfrak{G}_{n,k,E,S_n^*} f, g \rangle = \langle S_n^k \mathfrak{D}_{n,k,E,S_n^*} f, S_n^k \mathfrak{D}_{n,k,E,S_n^*} g \rangle = \sum_{j=0}^{\infty} \frac{\mu_{n,j}^2}{\mu_{n,j+k}} \langle f_j, g_j \rangle$$

from which we conclude that

$$\mathfrak{G}_{n,k,E,S_n^*}: \sum_{j=0}^{\infty} f_j z^j \rightarrow \sum_{j=0}^{\infty} \frac{\mu_{n,j}}{\mu_{n,j+k}} f_j z^j. \quad (5.20)$$

As  $\mu_{n,j}$  is decreasing in  $j$ , we see that  $\frac{\mu_{n,j}}{\mu_{n,j+k}} \geq 1$  and hence  $\mathfrak{G}_{n,k,E,S_n^*} \geq I_{\mathcal{A}_n(\mathcal{Y})}$  is strictly positive definite.  $\square$

## 6. FUNCTIONS $\Theta_{n,k}$ AND METRIC CONSTRAINTS

In this section we take a closer look at the transfer functions  $\Theta_{n,k}$  introduced by the realization formula (3.8). As was mentioned in Section 3, for an  $n$ -output stable pair  $(C, A)$ , the associated backward-shifted observability operators  $\mathfrak{D}_{n,k,C,A}$  are bounded for all  $k \geq 0$  as operators from  $\mathcal{X}$  into  $\mathcal{A}_n(\mathcal{Y})$ . In this case, the multiplication operator  $M_{\Theta_{n,k}}$  given (according to (3.8)) by

$$M_{\Theta_{n,k}} = \binom{k+n-1}{k} D_k + S_n \mathfrak{D}_{n,k+1,C,A} B_k: \mathcal{U}_k \rightarrow \mathcal{A}_n(\mathcal{Y}) \quad (6.1)$$

is also bounded. Therefore, the output function  $\widehat{y}$  in (3.6),

$$\widehat{y}(z) = \mathcal{O}_{n,C,A}x + \sum_{k=0}^N z^k \Theta_{n,k}(z) u_k \quad (6.2)$$

belongs to  $\mathcal{A}_n(\mathcal{Y})$  for every choice of  $x \in \mathcal{X}$  and  $u_k \in \mathcal{U}_k$  for each  $N = 1, 2, \dots$ . We next impose some additional metric relations on  $\begin{bmatrix} A & B_k \\ C & D_k \end{bmatrix}$ , specifically one or more of the relations

$$A^* \mathfrak{G}_{n,k+1,C,A} B_k + \binom{n+k-1}{k} \cdot C^* D_k = 0, \quad (6.3)$$

$$B_k^* \mathfrak{G}_{n,k+1,C,A} B_k + \binom{n+k-1}{k} \cdot D_k^* D_k \leq I_{\mathcal{U}_k}, \quad (6.4)$$

$$B_k^* \mathfrak{G}_{n,k+1,C,A} B_k + \binom{n+k-1}{k} \cdot D_k^* D_k = I_{\mathcal{U}_k}, \quad (6.5)$$

and show how these lead to boundedness and orthogonality properties for the associated multiplication operator  $M_{\Theta_k}$ . Due to equality (4.30), it turns out that relations (6.3) and (6.4) are equivalent to the matrix inequality

$$\begin{bmatrix} A^* & C^* \\ B_k^* & D_k^* \end{bmatrix} \begin{bmatrix} \mathfrak{G}_{n,k+1,C,A} & 0 \\ 0 & \binom{n+k-1}{k} \cdot I_{\mathcal{Y}} \end{bmatrix} \begin{bmatrix} A & B_k \\ C & D_k \end{bmatrix} \leq \begin{bmatrix} \mathfrak{G}_{n,k,C,A} & 0 \\ 0 & I_{\mathcal{U}_k} \end{bmatrix}, \quad (6.6)$$

while the equalities (6.3) and (6.5) are equivalent to the matrix equality

$$\begin{bmatrix} A^* & C^* \\ B_k^* & D_k^* \end{bmatrix} \begin{bmatrix} \mathfrak{G}_{n,k+1,C,A} & 0 \\ 0 & \binom{n+k-1}{k} \cdot I_{\mathcal{Y}} \end{bmatrix} \begin{bmatrix} A & B_k \\ C & D_k \end{bmatrix} = \begin{bmatrix} \mathfrak{G}_{n,k,C,A} & 0 \\ 0 & I_{\mathcal{U}_k} \end{bmatrix}. \quad (6.7)$$

The two latter conditions are of metric nature; they express the contractivity or isometric property of the colligation operator  $\mathbf{U}_k = \begin{bmatrix} A & B_k \\ C & D_k \end{bmatrix}$  with respect to certain weights. Note that in the classical case  $n = 1$ , the formula (6.10) below amounts to the following well-known fact: if the colligation operator  $\mathbf{U} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$  is isometric and if we set  $\Theta(z) = D + zC(I - zA)^{-1}B$ , then

$$\frac{I - \Theta(\zeta)^* \Theta(z)}{1 - z\bar{\zeta}} = B^*(I - \bar{\zeta}A^*)^{-1}(I - zA)^{-1}B. \quad (6.8)$$

**Lemma 6.1.** *Let  $(C, A)$  be an  $n$ -output stable pair and let  $\Theta_{n,k}$  be defined as in (3.8) for some integer  $k \geq 0$  and operators  $B_k \in \mathcal{L}(\mathcal{U}_k, \mathcal{X})$  and  $D_k \in \mathcal{L}(\mathcal{U}_k, \mathcal{Y})$ .*

(1) *If equality (6.3) holds, then*

- (a)  $\mathcal{O}_{n,C,A}x$  is orthogonal to  $S_n^k \Theta_{n,k}u$  for all  $x \in \mathcal{X}$  and  $u \in \mathcal{U}_k$ .
- (b)  $S_n^k \Theta_{n,k}u$  is orthogonal to  $S_n^m \Theta_{n,k}u'$  for all  $m > k$  and  $u, u' \in \mathcal{U}_k$ .

(2) *Moreover:*

- (a) *If inequality (6.4) holds, then the operator  $S_n^k M_{\Theta_{n,k}}$  is a contraction from  $\mathcal{U}_k$  into  $\mathcal{A}_n(\mathcal{Y})$ .*
- (b) *If both (6.3) and (6.4) hold, i.e., if (6.6) holds, then the operator  $S_n^k M_{\Theta_{n,k}}$  is a contraction from  $H^2(\mathcal{U}_k)$  into  $\mathcal{A}_n(\mathcal{Y})$ .*

(3) *Similarly:*

- (a) *If equality (6.5) holds, then the operator  $S_n^k M_{\Theta_{n,k}}$  is an isometry from  $\mathcal{U}_k$  into  $\mathcal{A}_n(\mathcal{Y})$ .*
- (b) *If (6.3) and (6.5) hold, i.e., if (6.7) holds, then*

$$\|S_n^k \Theta_{n,k}f\|_{\mathcal{A}_n(\mathcal{Y})}^2 = \|f\|_{H^2(\mathcal{U}_k)}^2 - \sum_{j=1}^{\infty} \|(I - S_n^* S_n)^{1/2} S_n^k \Theta_{n,k} S_1^{*j} f\|^2 \quad (6.9)$$

for every  $f \in H^2(\mathcal{U}_k)$ .

(c) If (6.7) holds, then

$$\begin{aligned} \frac{I_{\mathcal{U}_k}}{\mu_{n,k}} - \Theta_{n,k}(z)^* \Theta_{n,k}(\zeta) &= \mu_{n,k} B_k^* R_{n,k}(zA)^* \mathfrak{G}_{n,k+1,C,A} R_{n,k}(\zeta A) B_k \\ &\quad - z\bar{\zeta} \mu_{n,k} B_k^* R_{n,k+1}(zA)^* \mathfrak{G}_{n,k,C,A} R_{n,k+1}(\zeta A) B_k. \end{aligned} \quad (6.10)$$

*Proof of (1):* We first observe the power series expansion

$$\Theta_{n,k}(z) = \mu_{n,k}^{-1} D_k + \sum_{j=0}^{\infty} \mu_{n,j+k+1}^{-1} \cdot C A^j B_k z^{j+1}. \quad (6.11)$$

which is an immediate consequence of formulas (3.8) and (4.27). We then make use of expansions (3.7), (6.11) and the definition of the inner product in  $\mathcal{A}_n(\mathcal{Y})$  to get

$$\begin{aligned} \langle S_n^k \Theta_{n,k} u, \mathcal{O}_{n,C,A} x \rangle_{\mathcal{A}_n(\mathcal{Y})} &= \mu_{n,k} \cdot \left\langle \mu_{n,k}^{-1} D_k u, \mu_{n,k}^{-1} C A^k x \right\rangle_{\mathcal{Y}} \\ &\quad + \sum_{j=0}^{\infty} \mu_{n,j+k+1} \cdot \left\langle \mu_{n,j+k+1}^{-1} C A^j B_k u, \mu_{n,j+k+1}^{-1} C A^{j+k+1} x \right\rangle_{\mathcal{Y}} \\ &= \left\langle \left( \binom{n+k-1}{k} C^* D_k + A^* \left( \sum_{j=0}^{\infty} \binom{n+j+k}{j+k+1} A^{*j} C^* C A^j \right) B_k \right) u, A^k x \right\rangle_{\mathcal{X}} \\ &= \langle \left( \binom{n+k-1}{k} C^* D_k + A^* \mathfrak{G}_{n,k+1,C,A} B_k \right) u, A^k x \rangle_{\mathcal{X}} \text{ (by (4.28))} \\ &= 0 \text{ (by (6.3))}. \end{aligned}$$

which proves part (a). Verification of part (b) is quite similar: for  $m > k$  we have

$$\begin{aligned} \langle S_n^m \Theta_{n,k} u', S_n^k \Theta_{n,k} u \rangle_{\mathcal{A}_n(\mathcal{Y})} &= \mu_{n,m} \cdot \left\langle \mu_{n,k}^{-1} D_k u', \mu_{n,m}^{-1} C A^{m-k-1} B_k u \right\rangle_{\mathcal{Y}} \\ &\quad + \sum_{j=0}^{\infty} \mu_{n,j+m+1} \cdot \left\langle \mu_{n,j+k+1}^{-1} C A^j B_k u', \mu_{n,j+m+1}^{-1} C A^{j+m-k-1} B_k u \right\rangle_{\mathcal{Y}} \\ &= \left\langle \left( \binom{n+k-1}{k} C^* D_k + A^* \left( \sum_{j=0}^{\infty} \binom{n+j+k}{j+k+1} A^{*j} C^* C A^j \right) B_k \right) u', A^{m-k-1} B_k u \right\rangle_{\mathcal{X}} \\ &= \langle \left( \binom{n+k-1}{k} C^* D_k + A^* \mathfrak{G}_{n,k+1,C,A} B_k \right) u', A^{m-k-1} B_k u \rangle_{\mathcal{X}} = 0. \end{aligned}$$

□

*Proof of (2):* According to (6.4),

$$\binom{k+n-1}{k} \cdot \|D_k u\|_{\mathcal{Y}}^2 + \langle \mathfrak{G}_{n,k+1,C,A} B_k u, B_k u \rangle_{\mathcal{X}} \leq \|u\|_{\mathcal{U}_k}^2 \quad (6.12)$$

for all  $u \in \mathcal{U}_k$ . We now may make use of (5.18) to get

$$\begin{aligned} \|S_n^k \Theta_{n,k} u\|_{\mathcal{A}_n(\mathcal{Y})}^2 &= \| \left( \binom{k+n-1}{k} S_n^k D_k + S_n^{k+1} \mathfrak{D}_{n,k+1,C,A} B_k \right) u \|_{\mathcal{A}_n(\mathcal{Y})}^2 \\ &= \| \binom{k+n-1}{k} S_n^k D_k u \|_{\mathcal{A}_n(\mathcal{Y})}^2 + \| S_n^{k+1} \mathfrak{D}_{n,k+1,C,A} B_k u \|_{\mathcal{A}_n(\mathcal{Y})}^2 \\ &= \binom{k+n-1}{k} \|D_k u\|_{\mathcal{Y}}^2 + \langle \mathfrak{G}_{n,k+1,C,A} B_k u, B_k u \rangle_{\mathcal{X}} \leq \|u\|_{\mathcal{U}_k}^2. \end{aligned} \quad (6.13)$$

Thus,  $S_n^k M_{\Theta_{n,k}}$  is a contraction from  $\mathcal{U}_k$  to  $\mathcal{A}_n(\mathcal{Y})$  and part (a) follows. To prove part (b), following ideas from [29], under the assumption that both (6.3) and (6.4)

hold, we shall show that for any  $\mathcal{U}_k$ -valued polynomial  $f(z) = \sum_{j=0}^m f_j z^j$ ,

$$\|S_n^k \Theta_{n,k} f\|_{\mathcal{A}_n(\mathcal{Y})}^2 \leq \|f\|_{H^2(\mathcal{U}_k)}^2 = \sum_{j=0}^m \|f_j\|_{\mathcal{U}_k}^2. \quad (6.14)$$

Let  $S_1^*$  be the operator of backward shift on  $H^2(\mathcal{U}_k)$  so that for the polynomial  $f$  as above,  $(S_1^* f)(z) = \sum_{j=0}^{m-1} f_{j+1} z^j$ . By statements (1b) and (3a) of the lemma, we have

$$\begin{aligned} \|S_n^k \Theta_{n,k} f\|^2 &= \left\| \sum_{j=0}^m S_n^{k+j} \Theta_{n,k} f_j \right\|^2 \\ &= \|S_n^k \Theta_{n,k} f_0\|^2 + \left\| \sum_{j=1}^m S_n^{k+j} \Theta_{n,k} f_j \right\|^2 \quad (\text{by (1b)}) \\ &\leq \|f_0\|^2 + \left\| S_n^{k+1} \sum_{j=0}^{m-1} S_n^j \Theta_{n,k} f_{j+1} \right\|^2 \\ &= \|f_0\|^2 + \|S_n^{k+1} \Theta_{n,k} S_1^* f\|^2 \\ &= \|f_0\|^2 + \|S_n^k \Theta_{n,k} S_1^* f\|^2 - \left\| (I - S_n^* S_n)^{\frac{1}{2}} S_n^k \Theta_{n,k} S_1^* f \right\|^2. \end{aligned} \quad (6.15)$$

Replacing  $f$  by  $S_1^{*j} f$  in (6.15) gives

$$\begin{aligned} \|S_n^k \Theta_{n,k} S_1^{*j} f\|^2 &\leq \|f_j\|^2 + \|S_n^k \Theta_{n,k} (S_1^*)^{j+1} f\|^2 \\ &\quad - \left\| (I - S_n^* S_n)^{\frac{1}{2}} S_n^k \Theta_{n,k} (S_1^*)^{j+1} f \right\|^2 \end{aligned} \quad (6.16)$$

for  $j = 1, \dots, m$ . Iteration of the inequality (6.15) using (6.16) then gives

$$\begin{aligned} \|S_n^k \Theta_{n,k} f\|_{\mathcal{A}_n(\mathcal{Y})}^2 &\leq \sum_{j=0}^m \|f_j\|_{\mathcal{U}_k}^2 - \sum_{j=1}^m \left\| (I - S_n^* S_n)^{\frac{1}{2}} S_n^k \Theta_{n,k} S_1^{*j} f \right\|^2 \\ &\leq \sum_{j=0}^m \|f_j\|_{\mathcal{U}_k}^2. \end{aligned} \quad (6.17)$$

Letting  $m \rightarrow \infty$  in (6.17) now implies the validity of (6.14) for every  $f \in H^2(\mathcal{U}_k)$  and the proof of part (b) of (2) is now complete.  $\square$

*Proof of (3):* In case (6.5) holds, then (6.12) holds with equality and part (a) of (3) follows. If also (6.10) holds, then (6.15) holds with equality as well in (6.14), (6.16), and (6.17). Part (b) of (3) now follows by letting  $m \rightarrow \infty$  in (6.17).

It remains to verify the formula (6.10) under assumption (6.7). The identity (6.7) is equivalent to the collection of identities

$$\begin{aligned} A^* \mathfrak{G}_{n,k+1,C,A} A + \frac{C^* C}{\mu_{n,k}} &= \mathfrak{G}_{n,k,C,A}, \quad A^* \mathfrak{G}_{n,k+1,C,A} B_k + \frac{C^* D_k}{\mu_{n,k}} = 0, \\ B_k^* \mathfrak{G}_{n,k+1,C,A} B_k + \frac{D_k^* D_k}{\mu_{n,k}} &= I_{\mathcal{U}_k}. \end{aligned} \quad (6.18)$$

We use these relations to compute

$$\begin{aligned}
& \frac{I_{\mathcal{U}_k}}{\mu_{n,k}} - \Theta_{n,k}(z)^* \Theta_{n,k}(\zeta) = \\
& \frac{I_{\mathcal{U}_k}}{\mu_{n,k}} - \left[ \frac{D_k^*}{\mu_{n,k}} + \bar{z}B_k^*R_{n,k+1}(zA)^*C^* \right] \left[ \frac{D_k}{\mu_{n,k}} + \zeta CR_{n,k+1}(\zeta A)B_k \right] \\
& = \frac{I_{\mathcal{U}_k}}{\mu_{n,k}} - \frac{D_k^*D_k}{\mu_{n,k}^2} - \bar{z}B_k^*R_{n,k+1}(zA)^* \frac{C^*D_k}{\mu_{n,k}} - \zeta \frac{D_k^*C}{\mu_{n,k}} R_{n,k+1}(\zeta A)B_k \\
& \quad - \bar{z}\zeta B_k^*R_{n,k+1}(zA)^*C^*CR_{n,k+1}(\zeta A)B_k \\
& = \frac{1}{\mu_{n,k}} B_k^* \mathfrak{G}_{n,k+1,C,A} B_k + \bar{z}B_k^*R_{n,k+1}(zA)^* \mathfrak{G}_{n,k+1,C,A} B_k \\
& \quad + \zeta B_k^* \mathfrak{G}_{n,k+1,C,A} A R_{n,k+1}(\zeta A) B_k \\
& \quad - \mu_{n,k} \bar{z}\zeta B_k^*R_{n,k+1}(zA)^* (\mathfrak{G}_{n,k,C,A} - A^* \mathfrak{G}_{n,k+1,C,A} A) R_{n,k+1}(\zeta A) B_k
\end{aligned}$$

where we made use of (6.18) in the last step. By making use next of relation (2.14) we can continue the computation as

$$\begin{aligned}
& \frac{I_{\mathcal{U}_k}}{\mu_{n,k}} - \Theta_{n,k}(z)^* \Theta_{n,k}(\zeta) \\
& = \frac{1}{\mu_{n,k}} B_k^* \mathfrak{G}_{n,k+1,C,A} B_k + B_k^* \left( R_{n,k}(zA)^* - \frac{I}{\mu_{n,k}} \right) \mathfrak{G}_{n,k+1,C,A} B_k \\
& \quad + B_k^* \mathfrak{G}_{n,k+1,C,A} \left( R_{n,k}(\zeta A) - \frac{I}{\mu_{n,k}} \right) B_k \\
& \quad + \mu_{n,k} B_k^* \left( R_{n,k}(zA)^* - \frac{I}{\mu_{n,k}} \right) \mathfrak{G}_{n,k+1,C,A} \left( R_{n,k}(\zeta A) - \frac{I}{\mu_{n,k}} \right) B_k \\
& \quad - \mu_{n,k} \bar{z}\zeta B_k^*R_{n,k+1}(zA)^* \mathfrak{G}_{n,k,C,A} R_{n,k+1}(\zeta A) B_k \\
& = \mu_{n,k} B_k^* R_{n,k}(zA)^* \mathfrak{G}_{n,k+1,C,A} R_{n,k}(\zeta A) B_k \\
& \quad - \mu_{n,k} \bar{z}\zeta B_k^*R_{n,k+1}(zA)^* \mathfrak{G}_{n,k,C,A} R_{n,k+1}(\zeta A) B_k
\end{aligned}$$

verifying formula (6.10).  $\square$

The following result is an immediate consequence of Lemma 6.1.

**Corollary 6.2.** *Let us assume that the pair  $(C, A)$  is  $n$ -output stable and that relations (6.3), (6.4) hold for all  $k \geq 0$ . Then the representation (6.2) of the function  $\hat{y}$  is orthogonal in the metric of  $\mathcal{A}_n(\mathcal{Y})$  and*

$$\|\hat{y}\|_{\mathcal{A}_n(\mathcal{Y})}^2 = \|\mathcal{O}_{n,C,A}x\|^2 + \sum_{k=0}^{\infty} \|\Theta_{n,k}u_k\|^2 \leq \|\mathcal{G}_{n,C,A}^{\frac{1}{2}}x\|_{\mathcal{X}}^2 + \sum_{k=0}^{\infty} \|u_k\|_{\mathcal{U}_k}^2. \quad (6.19)$$

If relations (6.4) hold with equalities for all  $k \geq 0$ , then equality holds in (6.19).

Observe that in case the pair  $(C, A)$  is exactly  $n$ -observable (so that  $\mathfrak{G}_{n,k,C,A}$  is strictly positive definite for all  $k \geq 0$  by Proposition 5.7), the inequality (6.6) can equivalently be expressed as  $\|\Xi\| \leq 1$  where  $\Xi$  is the operator given by

$$\Xi := \begin{bmatrix} \mathfrak{G}_{n,k+1,C,A}^{1/2} & 0 \\ 0 & \frac{I_{\mathcal{Y}}}{\mu_{n,k}} \end{bmatrix} \begin{bmatrix} A & B_k \\ C & D_k \end{bmatrix} \begin{bmatrix} \mathfrak{G}_{n,k,C,A}^{-1/2} & 0 \\ 0 & I_{\mathcal{U}_k} \end{bmatrix} : \begin{bmatrix} \mathcal{X} \\ \mathcal{U}_k \end{bmatrix} \rightarrow \begin{bmatrix} \mathcal{X} \\ \mathcal{Y} \end{bmatrix}. \quad (6.20)$$

Another equivalent condition is that  $\|\Xi^*\| \leq 1$  which in turn can be expressed as

$$\begin{bmatrix} A & B_k \\ C & D_k \end{bmatrix} \begin{bmatrix} \mathfrak{G}_{n,k,C,A}^{-1} & 0 \\ 0 & I_{\mathcal{U}_k} \end{bmatrix} \begin{bmatrix} A^* & C^* \\ B_k^* & D_k^* \end{bmatrix} \leq \begin{bmatrix} \mathfrak{G}_{n,k+1,C,A}^{-1} & 0 \\ 0 & \mu_{n,k} I_{\mathcal{Y}} \end{bmatrix}. \quad (6.21)$$

Note that equality (6.7) means that the operator  $\Xi$  is isometric. Of particular interest is the case where  $\Xi$  is coisometric, i.e., where the colligation operator  $\mathbf{U}_k = \begin{bmatrix} A & B_k \\ C & D_k \end{bmatrix}$  is coisometric with respect to the weights indicated below:

$$\begin{bmatrix} A & B_k \\ C & D_k \end{bmatrix} \begin{bmatrix} \mathfrak{G}_{n,k,C,A}^{-1} & 0 \\ 0 & I_{\mathcal{U}_k} \end{bmatrix} \begin{bmatrix} A^* & C^* \\ B_k^* & D_k^* \end{bmatrix} = \begin{bmatrix} \mathfrak{G}_{n,k+1,C,A}^{-1} & 0 \\ 0 & \mu_{n,k} I_{\mathcal{Y}} \end{bmatrix}. \quad (6.22)$$

We can now derive the shifted weighted Bergman-space analogue of the familiar identity for the de Branges-Rovnyak kernel which is a dual version of (6.8): if  $\mathbf{U} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$  is coisometric and if we set  $\Theta(z) = D + zC(I - zA)^{-1}B$ , then

$$\frac{I - \Theta(z)\Theta(\zeta)^*}{1 - z\bar{\zeta}} = C(I - zA)^{-1}(I - \bar{\zeta}A^*)^{-1}C^*.$$

**Lemma 6.3.** *Let  $(C, A)$  be an exactly  $n$ -observable  $n$ -output stable pair and let  $\Theta_{n,k}$  be defined as in (3.8) for some operators  $B_k \in \mathcal{L}(\mathcal{U}_k, \mathcal{X})$  and  $D_k \in \mathcal{L}(\mathcal{U}_k)$  subject to equality (6.22). Then*

$$\begin{aligned} \frac{I_{\mathcal{Y}}}{\mu_{n,k}} - \Theta_{n,k}(z)\Theta_{n,k}(\zeta)^* &= CR_{n,k}(zA)\mathfrak{G}_{n,k,C,A}^{-1}R_{n,k}(\zeta A)^*C^*, \\ &\quad - z\bar{\zeta} \cdot CR_{n,k+1}(zA)\mathfrak{G}_{n,k+1,C,A}^{-1}R_{n,k+1}(\zeta A)^*C^*. \end{aligned} \quad (6.23)$$

*Proof.* The proof parallels the verification of the identity (6.10) done above. The weighted-coisometry condition (6.22) gives us the set of equations

$$\begin{aligned} A\mathfrak{G}_{n,k,C,A}^{-1}A^* + B_kB_k^* &= \mathfrak{G}_{n,k+1,C,A}^{-1}, \quad C\mathfrak{G}_{n,k,C,A}^{-1}A^* + D_kB_k^* = 0 \\ C\mathfrak{G}_{n,k,C,A}^{-1}C^* + D_kD_k^* &= \mu_{n,k} I_{\mathcal{Y}}. \end{aligned} \quad (6.24)$$

We then compute:

$$\begin{aligned} &\frac{I_{\mathcal{Y}}}{\mu_{n,k}} - \Theta_{n,k}(z)\Theta_{n,k}(\zeta)^* \\ &= \frac{I_{\mathcal{Y}}}{\mu_{n,k}} - \left[ \frac{D_k}{\mu_{n,k}} + zCR_{n,k+1}(zA)B_k \right] \left[ \frac{D_k^*}{\mu_{n,k}} + \bar{\zeta}B_k^*R_{n,k+1}(\zeta A)^*C^* \right] \\ &= \frac{I_{\mathcal{Y}}}{\mu_{n,k}} - \frac{D_kD_k^*}{\mu_{n,k}^2} - zCR_{n,k+1}(zA)\frac{B_kD_k^*}{\mu_{n,k}} - \bar{\zeta}\frac{D_kB_k^*}{\mu_{n,k}}R_{n,k+1}(\zeta A)^*C^* \\ &\quad - z\bar{\zeta} \cdot CR_{n,k+1}(zA)B_kB_k^*R_{n,k+1}(\zeta A)^*C^* \quad (\text{by (6.24)}) \\ &= \frac{C}{\mu_{n,k}}\mathfrak{G}_{n,k,C,A}^{-1}\frac{C^*}{\mu_{n,k}} + zCR_{n,k+1}(zA)A\mathfrak{G}_{n,k,C,A}^{-1}\frac{C^*}{\mu_{n,k}} \\ &\quad + \bar{\zeta}\frac{C}{\mu_{n,k}}\mathfrak{G}_{n,k,C,A}^{-1}A^*R_{n,k+1}(\zeta A)^*C^* \\ &\quad - z\bar{\zeta} \cdot CR_{n,k+1}(zA)\left[\mathfrak{G}_{n,k+1,C,A}^{-1} - A\mathfrak{G}_{n,k,C,A}^{-1}A^*\right]R_{n,k+1}(\zeta A)^*C^* \\ &= C\left(\frac{I_{\mathcal{X}}}{\mu_{n,k}} + zR_{n,k+1}(zA)A\right)\mathfrak{G}_{n,k,C,A}^{-1}\left(\frac{I_{\mathcal{X}}}{\mu_{n,k}} + \bar{\zeta}A^*R_{n,k+1}(\zeta A)^*\right)C^* \\ &\quad - z\bar{\zeta} \cdot CR_{n,k+1}(zA)\mathfrak{G}_{n,k+1,C,A}^{-1}R_{n,k+1}(\zeta A)^*C^* \\ &= CR_{n,k}(zA)\mathfrak{G}_{n,k,C,A}^{-1}R_{n,k}(\zeta A)^*C^* \end{aligned}$$

$$-z\bar{\zeta} \cdot CR_{n,k+1}(zA)\mathfrak{G}_{n,k+1,C,A}^{-1}R_{n,k+1}(\zeta A)^*C^* \text{ (by (2.14))}. \quad \square$$

**Remark 6.4.** More generally, if  $\Theta_{n,k}(z)$  is given by (3.8) and if we do not assume the weighted coisometry condition (6.22), then the decomposition (6.23) holds in the more general form

$$\begin{aligned} \frac{I_{\mathcal{Y}}}{\mu_{n,k}} - \Theta_{n,k}(z)\Theta_{n,k}(\zeta)^* &= CR_{n,k}(zA)\mathfrak{G}_{n,k,C,A}^{-1}R_{n,k}(\zeta A)^*C^* \\ &\quad - z\bar{\zeta} \cdot CR_{n,k+1}(zA)\mathfrak{G}_{n,k+1,C,A}^{-1}R_{n,k+1}(\zeta A)^*C^* + \Xi_k(z, \zeta) \end{aligned}$$

where the defect kernel  $\Xi_k(z, \zeta)$  is given by

$$\begin{aligned} \Xi_k(z, \zeta) &= \left[ zCR_{n,k}(zA) \quad \frac{I_{\mathcal{Y}}}{\mu_{n,k}} \right] \left( \begin{bmatrix} \mathfrak{G}_{n,k+1}^{-1} & 0 \\ 0 & \mu_{n,k}I_{\mathcal{Y}} \end{bmatrix} \right. \\ &\quad \left. - \begin{bmatrix} A & B_k \\ C & D_k \end{bmatrix} \begin{bmatrix} \mathfrak{G}_{n,k}^{-1} & 0 \\ 0 & I_{\mathcal{Y}} \end{bmatrix} \begin{bmatrix} A^* & C^* \\ B_k^* & D_k^* \end{bmatrix} \right) \cdot \begin{bmatrix} \bar{\zeta}R_{n,k}(\zeta A)^*C^* \\ \frac{I_{\mathcal{Y}}}{\mu_{n,k}} \end{bmatrix}. \end{aligned}$$

Since equality (6.22) implies inequality (6.6), it follows that under assumption of Lemma 6.3, all the conclusions of parts (1) and (2) in Lemma 6.1 are true. To have all conclusions true, we need the operator (6.20) to be unitary.

**Lemma 6.5.** Suppose that we are given an integer  $k \geq 0$  and an exactly  $n$ -observable  $n$ -output-stable pair  $(C, A)$  with  $A \in \mathcal{L}(\mathcal{X})$  and  $C \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ . Then there exist operators  $B_k \in \mathcal{L}(\mathcal{U}_k, \mathcal{X})$  and  $D_k \in \mathcal{L}(\mathcal{U}_k, \mathcal{Y})$  such that equalities (6.22) and (6.6) hold. Explicitly, such  $B_k$  and  $C_k$  are essentially uniquely determined by solving the Cholesky factorization problem:

$$\begin{bmatrix} B_k \\ D_k \end{bmatrix} \begin{bmatrix} B_k^* & D_k^* \end{bmatrix} = \begin{bmatrix} \mathfrak{G}_{n,k+1,C,A}^{-1} & 0 \\ 0 & \mu_{n,k}I_{\mathcal{Y}} \end{bmatrix} - \begin{bmatrix} A \\ C \end{bmatrix} \mathfrak{G}_{n,k,C,A}^{-1} \begin{bmatrix} A^* & C^* \end{bmatrix} \quad (6.25)$$

subject to the additional constraint that the coefficient space  $\mathcal{U}_k$  be chosen so that  $\begin{bmatrix} B_k \\ D_k \end{bmatrix} : \mathcal{U}_k \rightarrow \mathcal{X} \oplus \mathcal{Y}$  is injective.

*Proof.* By Proposition 4.11, the weighted Stein identity (4.30) holds for each  $k \geq 1$ . Since  $(C, A)$  is exactly observable, the gramian  $\mathfrak{G}_{n,k,C,A}$  is strictly positive definite and then it follows from (4.30) that the operator

$$\begin{bmatrix} \mathfrak{G}_{n,k+1,C,A}^{\frac{1}{2}} A \mathfrak{G}_{n,k,C,A}^{-\frac{1}{2}} \\ \mu_{n,k}^{-\frac{1}{2}} C \mathfrak{G}_{n,k,C,A}^{-\frac{1}{2}} \end{bmatrix} : \mathcal{X} \rightarrow \mathcal{X} \oplus \mathcal{Y}$$

is an isometry. By extending this operator to a coisometric operator (6.20) we arrive at  $B_k$  and  $D_k$  solving (6.25). Further, extension of this operator to a unitary amounts to the additional restriction that  $\begin{bmatrix} B_k \\ D_k \end{bmatrix}$  be injective.  $\square$

## 7. BEURLING-LAX THEOREM FOR $\mathcal{A}_n$

We now present our four approaches to Beurling-Lax representations for the Bergman space setting; these were sketched for the classical case in the Introduction above.

We start with a general observation. If the subspace  $\mathcal{M} \subset \mathcal{A}_n(\mathcal{Y})$  is  $S_n$ -invariant, then  $\mathcal{M}^\perp$  is  $S_n^*$ -invariant; hence we may apply Theorem 5.1 part (3) (with  $\mathcal{M}^\perp$  in place of  $\mathcal{M}$  viewed as sitting isometrically inside  $\mathcal{A}_n(\mathcal{Y})$ ) to conclude that there is an

$n$ -isometric pair  $(C, A)$  so that  $\mathcal{M}^\perp = \text{Ran } \mathcal{O}_{n,C,A}$ ; in fact, as indicated in Theorem 5.1, we may take  $(C, A)$  to be the model output pair  $(C, A) = (E|_{\mathcal{M}^\perp}, S_n^*|_{\mathcal{M}^\perp})$ , and  $\mathcal{O}_{n,C,A}$  amounts to the inclusion map of  $\mathcal{M}^\perp$  into  $\mathcal{A}_n(\mathcal{Y})$ . Since  $\|\mathcal{O}_{n,C,A}x\|_{\mathcal{A}_n(\mathcal{Y})}^2 = \langle \mathcal{G}_{n,C,A}x, x \rangle_{\mathcal{X}}$  for every  $x \in \mathcal{X}$ , it follows that  $\mathcal{M}^\perp$  is a reproducing kernel Hilbert space with reproducing kernel

$$k_{\mathcal{M}^\perp}(z, \zeta) = C(I - zA)^{-n} \mathcal{G}_{n,C,A}^{-1}(I - \bar{\zeta}A^*)^{-n} C^*.$$

It then follows that  $\mathcal{M} = (\mathcal{M}^\perp)^\perp$  has reproducing kernel

$$k_{\mathcal{M}}(z, \zeta) = \frac{I_{\mathcal{Y}}}{(1 - z\bar{\zeta})^n} - C(I - zA)^{-n} \mathcal{G}_{n,C,A}^{-1}(I - \bar{\zeta}A^*)^{-n} C^*. \quad (7.1)$$

**7.1. The first approach: partially isometric multipliers.** Suppose that we are given a shift-invariant subspace  $\mathcal{M}$  contained in the Hardy space  $H^2(\mathcal{Y})$ . As we just observed in the previous paragraph, then  $\mathcal{M}^\perp$  is backward-shift-invariant and hence can be represented as the range  $\text{Ran } \mathcal{O}_{C,A}$  of an observable output pair  $(C, A)$ ; we can even take  $(C, A)$  to be the model output pair:  $C: f \mapsto f(0)$  and  $A$  equal to the backward shift restricted to  $\mathcal{X} = \mathcal{M}^\perp$ . From the identity

$$\mathcal{G}_{1,C,A} - A^* \mathcal{G}_{1,C,A} A = C^* C \quad (7.2)$$

(where  $\mathcal{G}_{1,C,A}$  is the identity operator in the model case), we see that the column matrix  $\begin{bmatrix} \mathcal{G}_{1,C,A}^{1/2} A \mathcal{G}_{1,C,A}^{-1/2} \\ C \mathcal{G}_{1,C,A}^{-1/2} \end{bmatrix}$  is isometric. Hence there exist operators  $\mathbf{B}_1: \tilde{\mathcal{U}}_1 \rightarrow \mathcal{X}$  ( $= \mathcal{M}^\perp$ ) and  $\mathbf{D}_1: \tilde{\mathcal{U}}_1 \rightarrow \mathcal{Y}$  (where  $\tilde{\mathcal{U}}_1$  is an appropriate coefficient Hilbert space) so that the operator matrix

$$\begin{bmatrix} \mathcal{G}_{1,C,A}^{1/2} A \mathcal{G}_{1,C,A}^{-1/2} & \mathcal{G}_{1,C,A}^{1/2} \mathbf{B}_1 \\ C \mathcal{G}_{1,C,A}^{-1/2} & \mathbf{D}_1 \end{bmatrix}: \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \end{bmatrix} \rightarrow \begin{bmatrix} \mathcal{X} \\ \mathcal{Y} \end{bmatrix} \quad (7.3)$$

is unitary. We then set

$$\Psi_1(z) = \mathbf{D}_1 + zC(I - zA)^{-1}\mathbf{B}_1. \quad (7.4)$$

Making use of the unitary property of the matrix (7.3), we deduce that  $\Psi_0$  is inner in the sense that its boundary values exist almost everywhere on  $\mathbb{T}$  and are coisometric, and that the kernel identity

$$\frac{I_{\mathcal{Y}} - \Psi_1(z)\Psi_1(\zeta)^*}{1 - z\bar{\zeta}} = C(I - zA)^{-1} \mathcal{G}_{1,C,A}^{-1}(I - \bar{\zeta}A^*)^{-1} C^* \quad (7.5)$$

holds. For the case where  $n = 1$ , the latter identity implies the factorization of the kernel (7.1)

$$k_{\mathcal{M}}(z, \zeta) = \frac{\Psi_1(z)\Psi_1(\zeta)^*}{1 - z\bar{\zeta}}. \quad (7.6)$$

Standard reproducing-kernel arguments (details appear in the proof of Theorem 7.1 below for the more general representation (7.8)) then imply that  $\mathcal{M} = \Psi_1 \cdot \mathcal{A}_1(\mathcal{U}) = \Psi_1 \cdot H^2(\mathcal{U})$ , i.e.,  $\Psi_1$  serves as a Beurling-Lax representer for  $\mathcal{M}$ . One can view this approach as a state-space adaptation of the de Branges-Rovnyak reproducing kernel approach (see [12, 13]); it has proved to work well in a number of multivariable settings (see e.g. [7, 8]). The next theorem presents a Bergman-space analogue of the representation (7.6).

**Theorem 7.1.** *Let  $\mathcal{M}$  be a closed  $S_n$ -invariant subspace of  $\mathcal{A}_n(\mathcal{Y})$ . Then there is a partially isometric multiplier*

$$M_F = [M_{F_1} \quad \cdots \quad M_{F_n}] : \bigoplus_{j=1}^n \mathcal{A}_j(\mathcal{U}_j) \rightarrow \mathcal{A}_n(\mathcal{Y})$$

(so  $F(z)$  has the form  $F(z) = [F_1(z) \quad \cdots \quad F_n(z)]$  with each  $F_j$  holomorphic on  $\mathbb{D}$  with values in  $\mathcal{L}(\mathcal{U}_j, \mathcal{Y})$  for appropriate coefficient Hilbert spaces  $\mathcal{U}_j$ ,  $(j = 1, \dots, n)$ , so that

$$\mathcal{M} = M_F \left( \bigoplus_{j=1}^n \mathcal{A}_j(\mathcal{U}_j) \right). \quad (7.7)$$

*Proof.* To prove the theorem, it suffices to produce coefficient Hilbert spaces  $\mathcal{U}_1, \dots, \mathcal{U}_n$  and holomorphic operator-valued functions  $F_j : \mathbb{D} \rightarrow \mathcal{L}(\mathcal{U}_j, \mathcal{Y})$  so that

$$k_{\mathcal{M}}(z, \zeta) = \sum_{j=1}^n \frac{F_j(z)F_j(\zeta)^*}{(1 - z\bar{\zeta})^j}, \quad (7.8)$$

where  $k_{\mathcal{M}}$  is given in (7.1). Indeed, assume that (7.8) holds. Recall that  $\mathcal{A}_j$  is the reproducing kernel Hilbert space with reproducing kernel  $\mathbf{k}_j(z, \zeta) = (1 - z\bar{\zeta})^{-j}$  and define the operator  $X : \mathcal{M} = \mathcal{H}(k_{\mathcal{M}}) \rightarrow \bigoplus_{\ell=1}^n \mathcal{A}_{\ell}(\mathcal{U}_{\ell})$  by

$$X : k_{\mathcal{M}}(\cdot, \zeta)y \mapsto \begin{bmatrix} \mathbf{k}_1(\cdot, \zeta)F_1(\zeta)^*y \\ \vdots \\ \mathbf{k}_n(\cdot, \zeta)F_n(\zeta)^*y \end{bmatrix} \quad (y \in \mathcal{Y}, \zeta \in \mathbb{D}).$$

Then we have for  $y_i \in \mathcal{Y}$  and  $\zeta_i \in \mathbb{D}$ ,

$$\begin{aligned} \left\| X \left( \sum_{j=1}^N k_{\mathcal{M}}(\cdot, \zeta_j)y_j \right) \right\|^2 &= \left\langle \sum_{j=1}^N \begin{bmatrix} \mathbf{k}_1(\cdot, \zeta_j)F_1(\zeta_j)^*y_j \\ \vdots \\ \mathbf{k}_n(\cdot, \zeta_j)F_n(\zeta_j)^*y_j \end{bmatrix}, \sum_{i=1}^N \begin{bmatrix} \mathbf{k}_1(\cdot, \zeta_i)F_1(\zeta_i)^*y_i \\ \vdots \\ \mathbf{k}_n(\cdot, \zeta_i)F_n(\zeta_i)^*y_i \end{bmatrix} \right\rangle \\ &= \sum_{i,j=1}^N \left\langle \sum_{\ell=1}^N F_{\ell}(\zeta_i)\mathbf{k}_{\ell}(\zeta_i, \zeta_j)F_{\ell}(\zeta_j)^*y_j, y_i \right\rangle_{\mathcal{Y}} \\ &= \sum_{i,j=1}^N \langle k_{\mathcal{M}}(\zeta_i, \zeta_j)y_j, y_i \rangle_{\mathcal{Y}} = \left\| \sum_{j=1}^N k_{\mathcal{M}}(\cdot, \zeta_j)y_j \right\|_{\mathcal{M}}^2. \end{aligned}$$

We conclude that  $X$  extends uniquely from the span of the kernel functions  $k_{\mathcal{M}}(\cdot, \zeta)y$  ( $\zeta \in \mathbb{D}, y \in \mathcal{Y}$ ) to define an isometry from  $\mathcal{H}(k_{\mathcal{M}}) = \mathcal{M}$  into the direct sum space

$\bigoplus_{\ell=1}^n \mathcal{A}_\ell(\mathcal{U}_\ell)$ . Moreover, we compute, for  $h_\ell \in \mathcal{A}_\ell(\mathcal{U}_\ell)$  ( $\ell = 1, \dots, n$ ),

$$\begin{aligned} \left\langle \left( X^* \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \right) (\zeta), y \right\rangle_{\mathcal{Y}} &= \left\langle X^* \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix}, k_{\mathcal{M}}(\cdot, \zeta)y \right\rangle_{\mathcal{M}} \\ &= \left\langle \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix}, Xk_{\mathcal{M}}(\cdot, \zeta)y \right\rangle_{\bigoplus_{\ell=1}^n \mathcal{A}_\ell(\mathcal{U}_\ell)} \\ &= \left\langle \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix}, \begin{bmatrix} \mathbf{k}_1(\cdot, \zeta)F_1(\zeta)^*y \\ \vdots \\ \mathbf{k}_n(\cdot, \zeta)F_n(\zeta)^*y \end{bmatrix} \right\rangle_{\bigoplus_{\ell=1}^n \mathcal{A}_\ell(\mathcal{U}_\ell)} \\ &= \langle F_1(\zeta)h_1(\zeta) + \cdots + F_n(\zeta)h_n(\zeta), y \rangle_{\mathcal{Y}} \end{aligned}$$

from which we conclude that

$$X^* = M_F: \begin{bmatrix} h_1(z) \\ \vdots \\ h_n(z) \end{bmatrix} \mapsto F_1(z)h_1(z) + \cdots + F_n(z)h_n(z)$$

is the multiplication operator  $M_F = [M_{F_1} \ \cdots \ M_{F_n}]$  induced by the matrix function  $F(z) = [F_1(z) \ \cdots \ F_n(z)]$  and Theorem 7.1 then follows.

Thus the proof of Theorem 7.1 is reduced to the construction of holomorphic operator-valued functions  $F_1(z), \dots, F_n(z)$  so that (7.8) holds. The construction proceeds via an iterated unitary completion process as follows.

For  $j = 2, \dots, n$  we do a construction similar to that in (7.2)–(7.4) but based on the identity (see (4.10))

$$\mathcal{G}_{j,C,A} - A^* \mathcal{G}_{j,C,A} A = \mathcal{G}_{j-1,C,A} \quad (j = 2, \dots, n). \quad (7.9)$$

We find operators  $\mathbf{B}_j: \mathcal{U}_{n+1-j} \rightarrow \mathcal{X}$  and  $\mathbf{D}_j: \mathcal{U}_{n+1-j} \rightarrow \mathcal{Y}$  so that

$$\begin{bmatrix} A & \mathbf{B}_j \\ \mathcal{G}_{j-1,C,A}^{\frac{1}{2}} & \mathbf{D}_j \end{bmatrix} \begin{bmatrix} \mathcal{G}_{j,C,A}^{-1} & 0 \\ 0 & I_{\mathcal{U}_{n+1-j}} \end{bmatrix} \begin{bmatrix} A^* & \mathcal{G}_{j-1,C,A}^{\frac{1}{2}} \\ \mathbf{B}_j^* & \mathbf{D}_j^* \end{bmatrix} = \begin{bmatrix} \mathcal{G}_{j,C,A}^{-1} & 0 \\ 0 & I_{\mathcal{Y}} \end{bmatrix} \quad (7.10)$$

and

$$\begin{bmatrix} A^* & \mathcal{G}_{j-1,C,A}^{\frac{1}{2}} \\ \mathbf{B}_j^* & \mathbf{D}_j^* \end{bmatrix} \begin{bmatrix} \mathcal{G}_{j,C,A} & 0 \\ 0 & I_{\mathcal{Y}} \end{bmatrix} \begin{bmatrix} A & \mathbf{B}_j \\ \mathcal{G}_{j-1,C,A}^{\frac{1}{2}} & \mathbf{D}_j \end{bmatrix} = \begin{bmatrix} \mathcal{G}_{j,C,A} & 0 \\ 0 & I_{\mathcal{U}_{n+1-j}} \end{bmatrix}. \quad (7.11)$$

In fact, the latter equalities determine  $\mathbf{B}_j$  and  $\mathbf{D}_j$  uniquely up to a common unitary factor on the right:

$$\mathbf{B}_j = (\mathcal{G}_{j,C,A}^{-1} - A\mathcal{G}_{j,C,A}^{-1}A^*)^{\frac{1}{2}}, \quad \mathbf{D}_j = -\mathcal{G}_{j-1,C,A}^{-\frac{1}{2}}A^*\mathcal{G}_{j,C,A}\mathbf{B}_j. \quad (7.12)$$

We now define the functions

$$\Psi_j(z) = \mathbf{D}_j + z\mathcal{G}_{j-1,C,A}^{\frac{1}{2}}(I - zA)^{-1}\mathbf{B}_j \quad (7.13)$$

for  $j = 2, \dots, n$ , which are inner and satisfy the identities

$$\frac{I_{\mathcal{Y}} - \Psi_j(z)\Psi_j(\zeta)^*}{1 - z\bar{\zeta}} = \mathcal{G}_{j-1,C,A}^{\frac{1}{2}}(I - zA)^{-1}\mathcal{G}_{j,C,A}^{-1}(I - \bar{\zeta}A^*)^{-1}\mathcal{G}_{j-1,C,A}^{\frac{1}{2}}. \quad (7.14)$$

We finally define operator-valued functions  $F_\ell(z): \mathcal{U}_\ell \rightarrow \mathcal{Y}$  via

$$F_\ell(z) = C(I - zA)^{-(n-\ell)}\mathcal{G}_{n-\ell,C,A}^{-\frac{1}{2}}\Psi_{n+1-\ell}(z) \quad \text{for } \ell = 1, 2, \dots, n-1,$$

$$F_n(z) = \Psi_1(z) = \mathbf{D}_1 + zC(I - zA)^{-1}\mathbf{B}_1 \quad (7.15)$$

and claim that this choice of  $F_1, \dots, F_n$  satisfies (7.8) (with coefficient spaces  $\mathcal{U}_j = \mathcal{X}$  for  $j = 1, \dots, n - 1$ ). Indeed, substituting equality (7.14) (for  $j = n$ ) into the formula (7.1) for  $k_{\mathcal{M}}$  and then making use of formula (7.15) for  $F_1$ , we have

$$\begin{aligned} k_{\mathcal{M}}(z, \zeta) &= \frac{I_{\mathcal{Y}}}{(1 - z\bar{\zeta})^n} - C(I - zA)^{-(n-1)} \mathcal{G}_{n-1,C,A}^{-\frac{1}{2}} \left[ \frac{I - \Psi_n(z)\Psi_n(\zeta)^*}{1 - z\bar{\zeta}} \right] \\ &\quad \cdot \mathcal{G}_{n-1,C,A}^{-\frac{1}{2}} (I - \bar{\zeta}A^*)^{-(n-1)} C^* \\ &= \frac{I_{\mathcal{Y}}}{(1 - z\bar{\zeta})^n} + \frac{F_1(z)F_1(\zeta)^*}{1 - z\bar{\zeta}} \\ &\quad - C(I - zA)^{-(n-1)} \frac{\mathcal{G}_{n-1,C,A}^{-1}}{1 - z\bar{\zeta}} (I - \bar{\zeta}A^*)^{-(n-1)} C^*. \end{aligned}$$

One can now use an inductive argument to show that in general

$$\begin{aligned} k_{\mathcal{M}}(z, \zeta) &= \frac{I_{\mathcal{Y}}}{(1 - z\bar{\zeta})^n} + \sum_{\ell=1}^k \frac{F_{\ell}(z)F_{\ell}(\zeta)^*}{(1 - z\bar{\zeta})^{\ell}} \\ &\quad - C(I - zA)^{-(n-k)} \frac{\mathcal{G}_{n-k,C,A}^{-1}}{(1 - z\bar{\zeta})^k} (I - \bar{\zeta}A^*)^{-(n-k)} C^* \end{aligned} \quad (7.16)$$

for  $k = 1, 2, \dots, n - 1$ . For the final step, we note that the last term in (7.16) with  $k = n - 1$  is given by

$$\begin{aligned} C(I - zA)^{-1} \frac{\mathcal{G}_{1,C,A}^{-1}}{(1 - z\bar{\zeta})^{n-1}} (I - \bar{\zeta}A^*)^{-1} C^* &= \frac{1}{(1 - z\bar{\zeta})^{n-1}} \left[ \frac{I_{\mathcal{Y}} - \Psi_1(z)\Psi_1(\zeta)^*}{1 - z\bar{\zeta}} \right] \\ &= \frac{I_{\mathcal{Y}}}{(1 - z\bar{\zeta})^n} - \frac{F_n(z)F_n(\zeta)^*}{(1 - z\bar{\zeta})^n}. \end{aligned}$$

Plugging this expression into the  $k = n - 1$  case of (7.16) leaves us with (7.8) as wanted.  $\square$

There is an alternate formula for  $\Psi_j$  (and hence also for  $F_{\ell}$ ) which may prove useful in applications.

**Proposition 7.2.** *The function  $\Psi_j$  in (7.13) for  $j = 2, \dots, n$  can alternatively be given by*

$$\Psi_j(z) = \mathcal{G}_{j-1}^{\frac{1}{2}} (I - zA)^{-1} \mathcal{G}_j^{-1} (zI - A^*) \mathbf{B}_j^{-1}. \quad (7.17)$$

Hence  $F_{\ell}(z)$  can alternatively be given for  $\ell = 1, 2, \dots, n - 1$  by

$$\begin{aligned} F_{\ell}(z) &= C(I - zA)^{-(n-\ell)} \mathcal{G}_{n-\ell,C,A}^{-\frac{1}{2}} \Psi_{n+1-\ell}(z) \\ &= C(I - zA)^{-(n+1-\ell)} \mathcal{G}_{n+1-\ell}^{-1} (zI - A^*) \mathbf{B}_{n+1-\ell}^{-1}. \end{aligned} \quad (7.18)$$

*Proof.* In the proof, we shorten notation  $\mathcal{G}_{C,A,j}$  to  $\mathcal{G}_j$ . We substitute the second equality from (7.12) into (7.13) to get

$$\begin{aligned} \Psi_j(z) &= -\mathcal{G}_{j-1}^{-\frac{1}{2}} A^* \mathcal{G}_j B_j + z \mathcal{G}_{j-1}^{\frac{1}{2}} (I - zA)^{-1} \mathbf{B}_j \\ &= \mathcal{G}_{j-1}^{\frac{1}{2}} (I - zA)^{-1} [-(I - zA) \mathcal{G}_{j-1}^{-1} A^* \mathcal{G}_j + zI] \mathbf{B}_j \\ &= \mathcal{G}_{j-1}^{\frac{1}{2}} (I - zA)^{-1} [z(I + A \mathcal{G}_{j-1}^{-1} A^* \mathcal{G}_j) - \mathcal{G}_{j-1}^{-1} A^* \mathcal{G}_j] \mathbf{B}_j. \end{aligned} \quad (7.19)$$

We next observe the equalities

$$\mathcal{G}_{j-1}^{-1}A^*\mathcal{G}_j(I + A\mathcal{G}_{j-1}^{-1}A^*\mathcal{G}_j)^{-1} = \mathcal{G}_j^{-1}A^*\mathcal{G}_j,$$

and

$$\mathcal{G}_j + \mathcal{G}_j A \mathcal{G}_{j-1}^{-1} A^* \mathcal{G}_j = (\mathcal{G}_{C,A,j}^{-1} - A \mathcal{G}_{C,A,j}^{-1} A^*)^{-1},$$

which both follow from (7.9). We then use these equalities along with (7.19) and the first formula in (7.13) to get (7.17):

$$\begin{aligned}\Psi_j(z) &= \mathcal{G}_{j-1}^{\frac{1}{2}}(I - zA)^{-1}[zI - \mathcal{G}_j^{-1}A^*\mathcal{G}_j](I + A\mathcal{G}_{j-1}^{-1}A^*\mathcal{G}_j)\mathbf{B}_j \\ &= \mathcal{G}_{j-1}^{\frac{1}{2}}(I - zA)^{-1}\mathcal{G}_j^{-1}(zI - A^*)(\mathcal{G}_j + \mathcal{G}_j A \mathcal{G}_{j-1}^{-1} A^* \mathcal{G}_j)\mathbf{B}_j \\ &= \mathcal{G}_{j-1}^{\frac{1}{2}}(I - zA)^{-1}\mathcal{G}_j^{-1}(zI - A^*)(\mathcal{G}_j^{-1} - A \mathcal{G}_j^{-1} A^*)^{-1}\mathbf{B}_j \\ &= \mathcal{G}_{j-1}^{\frac{1}{2}}(I - zA)^{-1}\mathcal{G}_j^{-1}(zI - A^*)\mathbf{B}_j^{-1}.\end{aligned}$$

□

Given a shift invariant subspace  $\mathcal{M}$ , let us refer to the function  $F = [F_1 \ \cdots \ F_n]$  constructed as in Theorem 7.1 as the *partially isometric Bergman inner function* associated with  $\mathcal{M}$ .

**7.2. The second approach: shift-invariant subspaces contractively included in  $\mathcal{A}_n(\mathcal{Y})$ .** Let us say that the Hilbert space  $\mathcal{M}$  is contractively included in the Hilbert space  $\mathcal{H}$  if  $\mathcal{M} \subset \mathcal{H}$  as sets and moreover  $\|m\|_{\mathcal{H}} \leq \|m\|_{\mathcal{M}}$  for all  $m \in \mathcal{M}$ . We also say that an  $\mathcal{L}(\mathcal{U}, \mathcal{Y})$ -valued function  $\Theta$  is a *contractive multiplier* if the operator  $M_{\Theta} : f(z) \mapsto \Theta(z) \cdot f(z)$  of multiplication by  $\Theta$  defines a contractive operator from  $\mathcal{A}_n(\mathcal{U})$  to  $\mathcal{A}_n(\mathcal{Y})$ .

**Theorem 7.3.** *A Hilbert space  $\mathcal{M}$  is such that*

- (1)  $\mathcal{M}$  is contractively included in  $\mathcal{A}_n(\mathcal{Y})$ , i.e.  $\mathcal{M} \subset \mathcal{A}_n(\mathcal{Y})$  with  $\|f\|_{\mathcal{A}_n(\mathcal{Y})} \leq \|f\|_{\mathcal{M}}$  for all  $f \in \mathcal{M}$ ,
- (2)  $\mathcal{M}$  is  $S_n$ -invariant:  $S_n\mathcal{M} \subset \mathcal{M}$ ,
- (3) the operator  $A = (S_n|_{\mathcal{M}})^*$  is a strongly stable  $n$ -hypercontraction,

*if and only if there is a Hilbert space  $\mathcal{U}$  and a contractive multiplier  $\Theta$  so that*

$$\mathcal{M} = \Theta \cdot \mathcal{A}_n(\mathcal{U}) \tag{7.20}$$

*with lifted norm*

$$\|\Theta \cdot f\|_{\mathcal{M}} = \|Qf\|_{\mathcal{A}_n(\mathcal{U})} \tag{7.21}$$

*where  $Q$  is the orthogonal projection onto  $(\text{Ker } M_{\Theta})^{\perp}$ .*

*Proof.* We first verify sufficiency. Suppose that  $\mathcal{M}$  has the form (7.20) for a contractive multiplier  $\Theta$  with  $\mathcal{M}$ -norm given by (7.21). Since  $\|M_{\Theta}\| \leq 1$  it follows that

$$\|\Theta f\|_{\mathcal{A}_n(\mathcal{Y})} = \|M_{\Theta} Q f\|_{\mathcal{A}_n(\mathcal{Y})} \leq \|Q f\|_{\mathcal{A}_n(\mathcal{U})} = \|\Theta f\|_{\mathcal{M}}$$

i.e., (1) holds. From the intertwining equality  $S_n M_{\Theta} = M_{\Theta} S_n$ , property (2) follows. The latter intertwining equality also implies  $M_{\Theta} S_n|_{\text{Ker } M_{\Theta}} = 0$  which can be written equivalently in terms of the orthogonal projection  $Q$  onto  $(\text{Ker } M_{\Theta})^{\perp} \subset \mathcal{A}_n(\mathcal{U})$  as  $Q S_n(I - Q) = 0$ . Thus, we have

$$QS_n = QS_n Q \quad \text{and} \quad S_n^* Q = QS_n^* Q. \tag{7.22}$$

Furthermore, for every  $f, g \in \mathcal{A}_n(\mathcal{U})$ , we have

$$\begin{aligned} \langle \Theta g, A\Theta f \rangle_{\mathcal{M}} &= \langle S_n \Theta g, \Theta f \rangle_{\mathcal{M}} = \langle \Theta S_n g, \Theta f \rangle_{\mathcal{M}} = \langle Q S_n g, f \rangle_{\mathcal{A}_n(\mathcal{U})} \\ &= \langle Q S_n Q g, f \rangle_{\mathcal{A}_n(\mathcal{U})} = \langle Q g, S_n^* Q f \rangle_{\mathcal{A}_n(\mathcal{U})} = \langle \Theta g, \Theta S_n^* Q f \rangle_{\mathcal{M}}, \end{aligned}$$

which implies that  $A : \Theta f \rightarrow \Theta S_n^* Q f$ . Iterating the latter formula gives

$$A^j : \Theta f \rightarrow \Theta S_n^{*j} Q f \quad \text{for } j \geq 0. \quad (7.23)$$

Since  $S_n^*$  is a strongly stable  $n$ -hypercontraction on  $\mathcal{A}_n(\mathcal{Y})$ , we conclude from (7.22), (7.23), and (7.21) that

$$\begin{aligned} \lim_{N \rightarrow \infty} \|A^j \Theta f\|_{\mathcal{M}} &= \lim_{N \rightarrow \infty} \|\Theta S_n^{*j} Q f\|_{\mathcal{M}} \\ &= \lim_{N \rightarrow \infty} \|Q S_n^{*j} Q f\|_{\mathcal{A}_n(\mathcal{U})} = \lim_{N \rightarrow \infty} \|S_n^{*j} Q f\|_{\mathcal{A}_n(\mathcal{U})} = 0 \end{aligned}$$

and also

$$\sum_{j=0}^k (-1)^j \binom{k}{j} \|A^j \Theta f\|_{\mathcal{M}}^2 = \sum_{j=0}^k (-1)^j \binom{k}{j} \|S_n^{*j} Q f\|_{\mathcal{A}_n(\mathcal{U})}^2 \geq 0$$

for  $k = 1, \dots, n$  which shows that  $A$  is strongly stable  $n$ -hypercontraction on  $\mathcal{M}$  and therefore completes the proof of sufficiency.

Suppose now that the Hilbert space  $\mathcal{M}$  satisfies conditions (1), (2), (3) in the statement of the theorem. Using hypothesis (2) we can define the operator  $A := (S_n|_{\mathcal{M}})^*$  on  $\mathcal{M}$  and since it is  $n$ -hypercontractive by hypothesis (3), the operator  $\Gamma_{A,n}(I)$  is positive semidefinite. Choose the coefficient Hilbert space  $\mathcal{U}$  so that

$$\dim \mathcal{U} = \operatorname{rank} \Gamma_{A,n}(I)$$

and then choose the operator  $C : \mathcal{M} \rightarrow \mathcal{U}$  so that

$$C^* C = \Gamma_{n,A}(I) = \sum_{j=0}^n (-1)^j \binom{n}{j} A^{*j} A^j \geq 0.$$

Then  $(C, A)$  is an  $n$ -isometric pair and, since  $A$  is strongly stable by hypothesis (3). it follows from part (2) of Proposition 4.7 that the observability operator

$$\mathcal{O}_{n,C,A} : f \mapsto C(I - zA)^{-n} f$$

is an isometry from  $\mathcal{M}$  into  $\mathcal{A}_n(\mathcal{Y})$ . By part (1) of Theorem 5.1, we have the intertwining equality  $S_n^* \mathcal{O}_{n,C,A} = \mathcal{O}_{n,C,A} A$ . Taking adjoints then gives

$$(\mathcal{O}_{n,C,A})^* S_n = A^* (\mathcal{O}_{n,C,A})^*. \quad (7.24)$$

The inclusion map  $\iota : \mathcal{M} \rightarrow \mathcal{A}_n(\mathcal{Y})$  is a contraction by hypothesis (1). Moreover,  $\iota A^* = S_n \iota : \mathcal{M} \rightarrow \mathcal{A}_n(\mathcal{Y})$ . Therefore the operator

$$R = \iota \circ (\mathcal{O}_{n,C,A})^* : \mathcal{A}_n(\mathcal{U}) \rightarrow \mathcal{A}_n(\mathcal{Y})$$

is a contraction and

$$RS_n = \iota(\mathcal{O}_{n,C,A})^* S_n = \iota A^* (\mathcal{O}_{n,C,A})^* = S_n \iota(\mathcal{O}_{n,C,A})^* = S_n R.$$

Therefore (see [28])  $R$  is a multiplication operator, i.e., there is a contractive multiplier  $\Theta$  so that  $R = M_{\Theta}$ . Since  $\mathcal{O}_{n,C,A} : \mathcal{M} \rightarrow \mathcal{A}_n(\mathcal{Y})$  is an isometry, it follows that  $\operatorname{Ran}(\mathcal{O}_{n,C,A})^* = \mathcal{M}$  and also that  $\mathcal{M} = \Theta \cdot \mathcal{A}_n(\mathcal{U})$  with  $\mathcal{M}$ -norm given by (7.21). This completes the proof.  $\square$

**Remark 7.4.** It is of interest to consider Theorem 7.3 for the case where condition (1) is strengthened to

- (1')  $\mathcal{M}$  is isometrically contained in  $\mathcal{A}_n(\mathcal{Y})$ .

For the special case where  $n = 1$  (so  $\mathcal{A}_1(\mathcal{Y})$  is the Hardy space  $H^2(\mathcal{Y})$ ), it is not difficult to see that condition (3) is a consequence of (1') and (2). Then the representer  $\Theta$  is a coisometric multiplier and the proof of Theorem 7.3 reduces essentially to that of Rosenblum-Rovnyak (see [30]) for the case  $n = 1$ . The same approach has been adapted to the multivariable setting of the Drury-Arveson space in Arveson [6] and McCullough-Trent [24] as well as the freely noncommutative lift of the Drury-Arveson spaces to the Fock space [7, Theorem 2.14]. For the case  $n > 1$ , there is a new phenomenon: it is not the case that (1') and (2) imply (3): indeed, it is known that any Hilbert space operator can be recovered up to unitary equivalence as the compression of the Bergman shift  $S_2$  to the orthogonal difference  $\mathcal{M} \ominus \mathcal{N}$  of two nested  $S_2$ -invariant subspaces  $\mathcal{N} \subset \mathcal{M}$ ; in particular, the 2-hypercontractivity property of  $S_2^*$  is not preserved when one considers  $(S_2|_{\mathcal{M}})^*$  for an  $S_2$ -invariant subspace  $\mathcal{M}$ . A second manifestation of the inapplicability of Theorem 7.3 to general isometrically-included  $S_n$ -invariant subspaces of  $\mathcal{A}_n(\mathcal{Y})$  is given by Theorem 7.1:  $S_n$ -invariant subspaces of  $\mathcal{A}_n(\mathcal{Y})$  are modeled by a coisometric multiplier acting on a direct-sum Bergman space  $\bigoplus_{j=1}^n \mathcal{A}_j(\mathcal{U}_j)$  rather than just on a single Bergman space  $\mathcal{A}_n(\mathcal{U})$ .

**7.3. The third approach: Bergman-inner families.** Given an  $S_n$ -invariant subspace  $\mathcal{M}$  of  $\mathcal{A}_n(\mathcal{Y})$ , from the fact that

$$\bigcap_{k \geq 0} S_n^k \mathcal{M} \subset \bigcap_{k \geq 0} S_n^k \mathcal{A}(\mathcal{Y}) = \{0\}$$

one can see that we always have the orthogonal-sum decomposition for  $\mathcal{M}$ :

$$\mathcal{M} = \bigoplus_{k=0}^{\infty} (S_n^k \mathcal{M} \ominus S_n^{k+1} \mathcal{M}). \quad (7.25)$$

The space  $S_n^k \mathcal{M}$  has reproducing kernel  $(z, \zeta) \rightarrow z^k \bar{\zeta}^k k_{\mathcal{M}}(z, \zeta)$ , but since the operator  $S_n^k : \mathcal{A}_n(\mathcal{Y}) \rightarrow \mathcal{A}_n(\mathcal{Y})$  is only a contraction and not an isometry if  $n > 1$ , the metric induced by the above reproducing kernel is different from the metric of  $\mathcal{A}_n(\mathcal{Y})$ . To get the reproducing kernel for  $S_n^k \mathcal{M}$  consistent with the metric of  $\mathcal{A}_n(\mathcal{Y})$ , we need to consider backward-shifted version  $\mathfrak{D}_{n,k,C,A}$  of the observability operator  $\mathcal{O}_{n,C,A}$  and the backward-shifted version  $\mathfrak{G}_{n,k,C,A}$  of the observability gramian  $\mathcal{G}_{n,C,A}$  introduced in Section 4 above. It turns out that the space  $(S_n^k \mathcal{M})^\perp$  has a nice characterization in terms of the operator  $\mathfrak{D}_{n,k,C,A}$ .

**Proposition 7.5.** *The space  $(S_n^k \mathcal{M})^\perp$  is characterized as*

$$(S_n^k \mathcal{M})^\perp = \left( \bigoplus_{j=0}^{k-1} S_n^j \mathcal{Y} \right) \bigoplus S_n^k \text{Ran } \mathfrak{D}_{n,k,C,A} \quad (7.26)$$

where we identify the first term with the subspace of polynomials of degree at most  $k - 1$  in  $\mathcal{A}_n(\mathcal{Y})$ .

*Proof.* We wish to characterize all functions  $f(z) = \sum_{j=0}^{\infty} f_j z^j$  which are orthogonal to  $S_n^k \mathcal{M}$  in  $\mathcal{A}_n(\mathcal{Y})$ . We may write  $f(z) = p(z) + z^k \tilde{f}(z)$  where  $p(z) = \sum_{j=0}^{k-1} f_j z^j$  and  $\tilde{f} \in \mathcal{A}_n(\mathcal{Y})$ . Clearly polynomials of degree at most  $k-1$  are orthogonal to  $S_n^k \mathcal{M}$ , so it suffices to characterize which functions of the form  $z^k \tilde{f}(z)$  are orthogonal to  $S_n^k \mathcal{M}$ . To this end, observe that  $S_n^k \tilde{f}$  is orthogonal to  $S_n^k \mathcal{M}$  if and only if the function  $(S_n^k)^* S_n^k \tilde{f}$  belongs to  $\mathcal{M}^\perp = \text{Ran } \mathcal{O}_{n,C,A}$ . It follows from the formulas (5.2) and (1.12) that

$$(S_n^k)^* S_n^k : \sum_{j=0}^{\infty} \tilde{f}_j z^j \rightarrow \sum_{j=0}^{\infty} \frac{\mu_{n,j+k}}{\mu_{n,j}} \tilde{f}_j z^j.$$

We thus conclude that  $S_n^k \tilde{f}$  is orthogonal to  $S_n^k \mathcal{M}$  if and only if there exists a vector  $x \in \mathcal{X}$  such that

$$\sum_{j=0}^{\infty} \frac{\mu_{n,j+k}}{\mu_{n,j}} \tilde{f}_j z^j = C(I - zA)^{-n} x = \sum_{j=0}^{\infty} \left( \frac{1}{\mu_{n,j}} \cdot CA^j x \right) z^j.$$

Equating the corresponding Taylor coefficients gives

$$\tilde{f}_j = \frac{1}{\mu_{n,j+k}} \cdot CA^j x = \binom{n+j+k-1}{j+k} \cdot CA^j x \quad \text{for all } j \geq 0,$$

and therefore,

$$\tilde{f}(z) = \sum_{j=0}^{\infty} \tilde{f}_j z^j = \sum_{j=0}^{\infty} \left( \binom{n+j+k-1}{j+k} \cdot CA^j x \right) z^j = \mathfrak{D}_{n,k,C,A} x,$$

by (4.27). Thus,  $\tilde{f} \in \text{Ran } \mathfrak{D}_{n,k,C,A}$ . As the analysis is necessary and sufficient, the result follows.  $\square$

With characterization (7.26) in hand, it is straightforward to derive the kernel function for the space  $S_n^k \mathcal{M}$  with respect to the metric inherited from  $\mathcal{A}_n(\mathcal{Y})$ .

**Proposition 7.6.** *Let  $\mathcal{M}$  be a closed shift-invariant subspace of  $\mathcal{A}_n(\mathcal{Y})$  with reproducing kernel  $k_{\mathcal{M}}$  given by (7.1). Then the reproducing kernel functions for the closures of  $S_n^k \mathcal{M}$  and of  $S_n^k \mathcal{M} \ominus S_n^{k+1} \mathcal{M}$  in the metric of  $\mathcal{A}_n(\mathcal{Y})$  are given by*

$$k_{S_n^k \mathcal{M}}(z, \zeta) = z^k \bar{\zeta}^k \left( \sum_{\ell=1}^n \frac{\binom{\ell+k-2}{\ell-1} \cdot I_{\mathcal{Y}}}{(1 - z\bar{\zeta})^{n-\ell+1}} - CR_{n,k}(zA) \mathfrak{G}_{n,k,C,A}^{-1} R_{n,k}(\zeta A)^* C^* \right), \quad (7.27)$$

$$\begin{aligned} k_{S_n^k \mathcal{M} \ominus S_n^{k+1} \mathcal{M}}(z, \zeta) &= z^k \bar{\zeta}^k \left( \binom{n+k-1}{k} \cdot I_{\mathcal{Y}} - CR_{n,k}(zA) \mathfrak{G}_{n,k,C,A}^{-1} R_{n,k}(\zeta A)^* C^* \right. \\ &\quad \left. + z\bar{\zeta} CR_{n,k+1}(zA) \mathfrak{G}_{n,k+1,C,A}^{-1} R_{n,k+1}(\zeta A)^* C^* \right). \end{aligned} \quad (7.28)$$

*Proof.* By Theorem 5.6, the kernel  $\mathfrak{K}$  for the space  $S_n^k \text{Ran } \mathfrak{D}_{n,k,C,A}$  (with inner product induced by  $\mathcal{A}_n(\mathcal{Y})$ ) is given by (5.17). From the formula (7.26) for  $(S_n \mathcal{M})^\perp$

we deduce that

$$\begin{aligned} S_n^k \mathcal{M} &= \left( \bigoplus_{j=0}^{k-1} S_n^j \mathcal{Y} \right)^\perp \cap (S_n^k \operatorname{Ran} \mathfrak{D}_{n,k,C,A})^\perp \\ &= S_n^k \mathcal{A}_n(\mathcal{Y}) \ominus (S_n^k \operatorname{Ran} \mathfrak{D}_{n,k,C,A}). \end{aligned} \quad (7.29)$$

Since the reproducing kernel for the subspace  $S_n^j \mathcal{Y}$  of  $\mathcal{A}_n(\mathcal{Y})$  is  $z^j \bar{\zeta}^j \binom{n+j-1}{j} \cdot I_{\mathcal{Y}}$ ,

we deduce that  $S_n^k \mathcal{A}_n(\mathcal{Y}) = \mathcal{A}_n(\mathcal{Y}) \ominus \left( \bigoplus_{j=0}^{k-1} S_n^j \mathcal{Y} \right)$  has reproducing kernel

$$\begin{aligned} k_{S_n^k \mathcal{A}_n(\mathcal{Y})}(z, \zeta) &= \frac{I_{\mathcal{Y}}}{(1 - z\bar{\zeta})^n} - \sum_{j=0}^{k-1} \binom{n+j-1}{j} z^j \bar{\zeta}^j I_{\mathcal{Y}} \\ &= \sum_{j=k}^{\infty} \binom{n+j-1}{j} z^j \bar{\zeta}^j I_{\mathcal{Y}} = z^k \bar{\zeta}^k \left( \sum_{j=0}^{\infty} \binom{n+j+k-1}{j+k} z^j \bar{\zeta}^j I_{\mathcal{Y}} \right). \end{aligned}$$

If we now make use of the Chu-Vandermonde identity (2.6), we can continue the computation as follows:

$$\begin{aligned} k_{S_n^k \mathcal{A}_n(\mathcal{Y})}(z, \zeta) &= z^k \bar{\zeta}^k \left( \sum_{j=0}^{\infty} \sum_{\ell=1}^n \binom{\ell+k-2}{\ell-1} \cdot \binom{n+j-\ell}{j} z^j \bar{\zeta}^j \right) I_{\mathcal{Y}} \\ &= z^k \bar{\zeta}^k \sum_{\ell=1}^n \binom{\ell+k-2}{\ell-1} \left( \sum_{j=0}^{\infty} \binom{n+j-\ell}{j} z^j \bar{\zeta}^j \right) I_{\mathcal{Y}} \\ &= z^k \bar{\zeta}^k \sum_{\ell=1}^n \frac{\binom{\ell+k-2}{\ell-1} I_{\mathcal{Y}}}{(1 - z\bar{\zeta})^{n-\ell+1}}. \end{aligned} \quad (7.30)$$

From the representation for  $S_n^k \mathcal{M}$  in (7.29), we see that

$$k_{S_n^k \mathcal{M}}(z, \zeta) = k_{S_n^k \mathcal{A}_n(\mathcal{Y})}(z, \zeta) - \mathfrak{K}_k(z, \zeta) \quad (7.31)$$

where  $\mathfrak{K}_k$  is the reproducing kernel for  $S_n^k \operatorname{Ran} \mathfrak{D}_{n,k,C,A}$ . Combining this with (7.30) and the formula (5.17) for  $\mathfrak{K}_k$  leads to the formula (7.27) for  $k_{S_n^k \mathcal{M}}$ .

It remains to verify the formula (7.28). Toward this end, note that in view of (7.31) and (7.30),

$$\begin{aligned} k_{S_n^k \mathcal{M} \ominus S_n^{k+1} \mathcal{M}}(z, \zeta) &= k_{S_n^k \mathcal{M}}(z, \zeta) - k_{S_n^{k+1} \mathcal{M}}(z, \zeta) \\ &= k_{S_n^k \mathcal{A}_n(\mathcal{Y})}(z, \zeta) - k_{S_n^{k+1} \mathcal{A}_n(\mathcal{Y})}(z, \zeta) - \mathfrak{K}_k(z, \zeta) + \mathfrak{K}_{k+1}(z, \zeta) \\ &= k_{S_n^k \mathcal{Y}}(z, \zeta) - \mathfrak{K}_k(z, \zeta) + \mathfrak{K}_{k+1}(z, \zeta) \\ &= z^k \bar{\zeta}^k \binom{n+k-1}{k} \cdot I_{\mathcal{Y}} - \mathfrak{K}_k(z, \zeta) + \mathfrak{K}_{k+1}(z, \zeta). \end{aligned}$$

Plugging (5.17) into this last expression leads to (7.28) as wanted.  $\square$

**Lemma 7.7.** *Given an integer  $k \geq 1$  and an exactly observable  $n$ -output stable pair  $(C, A)$ , construct operators  $B_k \in \mathcal{L}(\mathcal{U}_k, \mathcal{X})$  and  $D_k \in \mathcal{L}(\mathcal{U}_k, \mathcal{Y})$  as in Lemma*

6.5 and let  $\Theta_{n,k}$  be the associated function given by (3.8). Then the kernel (7.28) can be factored as

$$k_{S_n^k \mathcal{M} \ominus S_n^{k+1} \mathcal{M}}(z, \zeta) = z^k \bar{\zeta}^k \Theta_{n,k}(z) \Theta_{n,k}(\zeta)^*. \quad (7.32)$$

*Proof.* By Lemma 6.3, identity (6.23) holds. Multiplying both parts of (6.23) by  $z^k \bar{\zeta}^k$  and combining the obtained equality with (7.28) we easily get (7.32).  $\square$

Let us say that a family of operator-valued functions  $\{\Theta_{n,k}: \mathbb{D} \rightarrow \mathcal{L}(\mathcal{U}_k, \mathcal{Y})\}_{k=0}^\infty$  is a *Bergman-inner family* if, for each  $k = 0, 1, 2, \dots$ , we have:

- (1)  $M_{\Theta_{n,k}}: S_1^k \mathcal{U}_k \rightarrow \mathcal{A}_n(\mathcal{Y})$  is isometric,
- (2)  $M_{\Theta_{n,k}}(S_1^k \mathcal{U}_k)$  is orthogonal (in  $\mathcal{A}_n(\mathcal{Y})$ ) to  $\overline{\text{span}}\{M_{\Theta_{n,\ell}} S_1^\ell \mathcal{U}_\ell: \ell > k\}$ , and
- (3)  $S_n^{k+1} M_{\Theta_{n,k}} \mathcal{U}_k \subset \overline{\text{span}}\{S_n^\ell M_{\Theta_{n,\ell}} \mathcal{U}_\ell: \ell > k\}$ .

If  $\{\Theta_{n,k}\}_{k \geq 0}$  is a Bergman-inner family and we set

$$\mathcal{M} = \bigoplus_{k=0}^{\infty} \Theta_{n,k} S_1^k \mathcal{U}_k \subset \mathcal{A}_n(\mathcal{Y}), \quad (7.33)$$

it then follows that the multiplication operator

$$M_\Theta = [M_{\Theta_0} \quad M_{\Theta_1} \quad M_{\Theta_2} \quad \cdots] : \bigoplus_{k=0}^{\infty} S_1^k \mathcal{U}_k \rightarrow \mathcal{A}_n(\mathcal{Y})$$

maps the *time-varying Hardy space*  $H^2(\{\mathcal{U}_k\}_{k \geq 0}) := \bigoplus_{k=0}^{\infty} S_1^k \mathcal{U}_k$  (where  $\mathbf{u} = \bigoplus_{k=0}^{\infty} z^k u_k \in H^2(\{\mathcal{U}_k\})$ ) is assigned the Hardy-space norm  $\|\mathbf{u}\|^2 = \sum_{k=0}^{\infty} \|u_k\|^2$ ) unitarily onto the

$S_n$ -invariant subspace  $\mathcal{M} \subset \mathcal{A}_n(\mathcal{Y})$ . Putting all the pieces together, we arrive at the following converse of all these observations which amounts to our third analogue of the Beurling-Lax theorem for the Bergman space setting.

**Theorem 7.8.** *Let  $\mathcal{M}$  be a closed  $S_n$ -invariant subspace of  $\mathcal{A}_n(\mathcal{Y})$ . Then there is a Bergman inner family  $\{\Theta_{n,k}\}_{k \geq 0}$  so that  $M_\Theta H^2(\{\mathcal{U}_k\}_{k \geq 0})$  (with  $M_\Theta$  as in (7.33)).*

*Furthermore, given the  $S_n$ -invariant subspace  $\mathcal{M} \subset \mathcal{A}_n(\mathcal{Y})$ , the associated representing Bergman-inner family  $\{\Theta_{n,k}\}_{k \geq 0}$  can be constructed according to the following algorithm:*

- (1) Set  $\mathcal{X} = \mathcal{M}^\perp$  and define  $A \in \mathcal{L}(\mathcal{X})$  and  $C \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$  by

$$A = S_n^*|_{\mathcal{M}^\perp}, \quad Cf = f(0) \quad \text{for } f \in \mathcal{M}^\perp.$$

- (2) Construct  $\begin{bmatrix} B_k \\ D_k \end{bmatrix}$  by solving the Cholesky factorization problem (6.25) in Lemma 6.5.

$$(3) \quad \text{Set } \Theta_{n,k}(z) = \frac{D_k}{\mu_{n,k}} + z C R_{n,k+1}(zA) B_k.$$

Alternatively, Bergman-inner families  $\{\Theta_{n,k}\}_{k \geq 0}$  can be constructed from an arbitrary strongly stable  $n$ -hypercontraction as follows. Let  $A \in \mathcal{L}(\mathcal{X})$  be any strongly stable  $n$ -hypercontraction on a Hilbert space  $\mathcal{X}$  and choose  $C \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$  with dense range so that

$$C^* C = (I - B_A)^n [I_{\mathcal{X}}]$$

(One such choice is  $C = D_{n,A}: \mathcal{X} \rightarrow \mathcal{Y}$  where  $D_{n,A} = (\sum_{k=0}^n (-1)^k \binom{n}{k} A^{*k} A^k)^{1/2}$  and  $\mathcal{Y} = \mathcal{D}_{n,A} := \overline{\text{Ran}} D_{n,A}$ .) Construct  $B_k$ ,  $D_k$  as in Step 2 above and then  $\Theta_{n,k}$

as in Step 3 above. Then  $\{\Theta_{n,k}\}_{k \geq 0}$  is a Bergman-inner family and any Bergman-inner family arises in this way.

**Example 7.9.** By way of an example, suppose that  $\mathcal{M} = \{0\} \subset \mathcal{A}_n(\mathcal{Y})$ , so  $\mathcal{X} := \mathcal{M}^\perp = \mathcal{A}_n(\mathcal{Y})$ . Then

$$A = S_n^*, \quad C = E: f \in \mathcal{A}_n(\mathcal{Y}) \mapsto f(0) \in \mathcal{Y}.$$

In the proof of part (3) of Proposition 5.7 (see (5.20)), we saw that

$$\mathfrak{G}_{n,k,E,S_n^*}: \sum_{j=0}^{\infty} f_j z^j \rightarrow \sum_{j=0}^{\infty} \frac{\mu_{n,j}}{\mu_{n,j+k}} f_j z^j. \quad (7.34)$$

and hence

$$\mathfrak{G}_{n,k,E,S_n^*}^{-1}: \sum_{j=0}^{\infty} f_j z^j \rightarrow \sum_{j=0}^{\infty} \frac{\mu_{n,j+k}}{\mu_{n,j}} f_j z^j. \quad (7.35)$$

One can then work out that

$$\mathfrak{G}_{n,k+1,E,S_n^*}^{-1} - S_n^* \mathfrak{G}_{n-,E,S_n^*}^{-1} S_n = 0, \quad \mu_{n,k} I_{\mathcal{Y}} - E \mathfrak{G}_{n,k,E,S_n^*}^{-1} E^* = 0.$$

Hence, solving for  $\begin{bmatrix} B_k \\ D_k \end{bmatrix}$  via the Cholesky factorization problem (6.25) (with  $(C, A) = (E, S_n^*)$ ) leads to  $\begin{bmatrix} B_k \\ D_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$  for all  $k$ , as is to be expected from the conclusion of Theorem 7.8 with  $\mathcal{M} = \{0\}$ .

**7.4. The fourth approach: wandering-subspace Bergman-inner functions.** Suppose that  $\mathcal{M} \subset \mathcal{A}_n(\mathcal{Y})$  is  $S_n$ -invariant. The subspace

$$\mathcal{E} := \mathcal{M} \ominus S\mathcal{M}$$

has the property that  $\mathcal{E} \subset \mathcal{M}$  and  $\mathcal{E} \perp S_n\mathcal{M}$ . If it is the case that  $\mathcal{E}$  generates  $\mathcal{M}$  in the sense that  $\mathcal{M} = \overline{\text{span}}\{S_n^k \mathcal{E}: k = 0, 1, 2, \dots\}$ , then one says that  $\mathcal{M}$  has the *wandering subspace property* (with wandering subspace equal to  $\mathcal{E}$ ). For the classical Hardy-space case  $n = 1$ , one has that  $\mathcal{E} = \Theta_0 \cdot \mathcal{U}_0$  for an appropriate coefficient Hilbert space  $\mathcal{U}_0$  and the inner function  $\Theta_0$  constructed in (7.4). Furthermore,  $\mathcal{M}$  admits an orthogonal decomposition

$$\mathcal{M} = \bigoplus_{j=0}^{\infty} S_1^j \mathcal{E} \quad (7.36)$$

and this decomposition coincides with that in (7.25).

If  $n > 1$ , the subspace  $S_n^j \mathcal{E}$  is orthogonal to  $\mathcal{E}$  for all  $j \geq 1$ , but it is not orthogonal to  $S_n^m \mathcal{E}$  if  $1 \leq m \neq j$ . Thus, the best one can expect is the equality

$$\mathcal{M} = \bigvee_{j=0}^{\infty} S_n^j \mathcal{E} \quad (7.37)$$

which indeed holds if  $n = 2, 3$ ; see [2], [31], [32] (see also [33] for a different approach). Letting  $k = 1$  in formula (7.28) and recalling that  $\mathfrak{G}_{n,0,C,A} = \mathcal{G}_{n,C,A}$ ,  $\mathfrak{D}_{n,0,C,A} = \mathcal{O}_{n,C,A}$  (see (4.29)) and  $\mu_{n,0} = 1$ , we conclude that the reproducing kernel for the subspace  $\mathcal{E}$  equals

$$\begin{aligned} k_{\mathcal{E}}(z, \zeta) &= I_{\mathcal{Y}} - C(I - zA)^{-n} \mathcal{G}_{n,C,A}^{-1} (I_{\mathcal{X}} - \bar{\zeta}A^*)^{-n} C^* \\ &\quad + z\bar{\zeta}C \left( \sum_{j=1}^n (I_{\mathcal{X}} - zA)^{-j} \right) \mathfrak{G}_{n,1,C,A}^{-1} \left( \sum_{j=1}^n (I_{\mathcal{X}} - \bar{\zeta}A^*)^{-j} \right) C^*. \end{aligned} \quad (7.38)$$

Following [27, 29] we say that the function  $\Theta$  is a *wandering-subspace Bergman-inner function* whenever

- (1)  $M_\Theta: \mathcal{U}_0 \rightarrow A_n(\mathcal{Y})$  is isometric, and
- (2)  $\Theta \cdot \mathcal{U}_0$  is orthogonal to  $S_n^\ell \Theta \cdot \mathcal{U}_0$  for  $\ell \geq 1$ .

In case  $M_\Theta \mathcal{U}_0 = \mathcal{E}$  is the wandering subspace for the  $S_n$ -invariant subspace  $\mathcal{M}$ , then we have the Beurling-Lax-type representation of  $\mathcal{M}$  as the closure of  $\Theta \cdot \mathbb{C}[z] \otimes \mathcal{U}_0$ , where  $\mathbb{C}[z]$  denotes the algebra of polynomials with coefficients in  $\mathbb{C}$ . Construction of a Beurling inner function for the wandering subspace  $\mathcal{E}$  of  $\mathcal{M}$  amounts to focusing on the first element  $\Theta_0$  in the Bergman-inner family associated with  $\mathcal{M}$ . Specifying Lemma 7.7 for the case  $k = 0$  then leads to the following.

**Theorem 7.10.** *Given a strongly stable  $n$ -output-pair  $(C, A)$ , then there exist operators  $B \in \mathcal{L}(\mathcal{U}, \mathcal{X})$  and  $D \in \mathcal{L}(\mathcal{U}, \mathcal{Y})$  which solve the Cholesky factorization problem*

$$\begin{bmatrix} B \\ D \end{bmatrix} \begin{bmatrix} B^* & D^* \end{bmatrix} = \begin{bmatrix} \mathfrak{G}_{n,1,C,A}^{-1} & 0 \\ 0 & I_{\mathcal{Y}} \end{bmatrix} - \begin{bmatrix} A \\ C \end{bmatrix} \mathcal{G}_{n,C,A}^{-1} \begin{bmatrix} A^* & C^* \end{bmatrix}. \quad (7.39)$$

Moreover, if  $\Theta$  is defined by

$$\Theta(z) = D + zC \sum_{j=1}^n (I - zA)^{-j} B \quad (7.40)$$

where  $\begin{bmatrix} B \\ D \end{bmatrix}$  solves (6.25) (with  $k = 1$ ), then:

- (1) The factorization  $k_{\mathcal{E}}(z, \zeta) = \Theta(z)\Theta(\zeta)^*$  holds and therefore, the multiplication operator  $M_\Theta$  maps  $\mathcal{U}$  onto  $\mathcal{E}$  unitarily.
- (2) The subspace  $\mathcal{E}$  is orthogonal to  $S_n^k \Theta_k \cdot \mathcal{U}$  for every  $k \geq 1$ .
- (3)  $M_\Theta$  is a contractive multiplier from the Hardy space  $H^2(\mathcal{U}_0)$  into the Bergman space  $\mathcal{A}_n(\mathcal{Y})$ . Hence, if  $\mathcal{M}$  has the wandering subspace property, we conclude that  $\mathcal{M}$  has the Beurling-Lax-type representation

$$\mathcal{M} = \mathcal{A}_n(\mathcal{Y})\text{-closure of } M_\Theta H^2(\mathcal{U}).$$

We note that the last statement in the theorem is a consequence of part (2b) of Lemma 6.1. We also note that realization formula (7.40) appears in [27, 29]) in slightly different terms.

**7.5. The case of zero-based shift-invariant subspaces with one zero.** In this concluding subsection we illustrate all four approaches by applying them to the simplest shift-invariant (zero-based) subspace

$$\mathcal{M} = \{f \in \mathcal{A}_n : f(\alpha) = 0\} \quad (7.41)$$

of  $\mathcal{A}_n$  where  $\alpha$  is a fixed point in  $\mathbb{D}$ . If  $n = 1$ , then

$$\mathcal{M} = b_\alpha(z)\mathcal{A}_1, \quad \text{where } b_\alpha(z) = \frac{z - \alpha}{1 - z\bar{\alpha}}.$$

If  $n > 1$  and  $\alpha = 0$ , the latter representation still holds. Let us assume that  $\alpha \neq 0$ . The one-dimensional space  $\mathcal{M}^\perp$  is spanned by the function  $\mathbf{k}_n(z, \alpha) = (1 - z\bar{\alpha})^{-n}$  and thus  $\mathcal{M}^\perp = \text{Ran } \mathcal{O}_{n,C,A}$  with  $A = \bar{\alpha}$  and any  $C \neq 0$ . We choose  $C$  to make the gramian  $\mathcal{G}_{n,C,A}$  equal one. Thus,

$$A = \bar{\alpha}, \quad C = (1 - |\alpha|^2)^{\frac{n}{2}}. \quad (7.42)$$

Then a simple calculation reveals that

$$\mathcal{G}_{j,C,A} = (1 - |\alpha|^2)^{n-j} \quad \text{for } j = 0, \dots, n. \quad (7.43)$$

An induction argument making use of the weighted Stein identity (4.30) combined with the initial condition  $\mathfrak{G}_{n,0,C,A} = \mathfrak{G}_{n,C,A} = 1$  then leads us to the formula for the shifted gramians:

$$\mathfrak{G}_{n,k,C,A} = \frac{1}{|\alpha|^{2k}} \cdot \left( 1 - (1 - |\alpha|^2)^n \cdot \sum_{j=0}^{k-1} \binom{n+j-1}{j} |\alpha|^{2j} \right) \quad \text{for } k \geq 1. \quad (7.44)$$

We may also relate the numbers (7.44) to the function  $R_{n,k}$  defined in (2.3):

$$\mathfrak{G}_{n,k,C,A} = (1 - |\alpha|^2)^n R_{n,k}(|\alpha|^2). \quad (7.45)$$

Formula (7.1) takes the form

$$k_{\mathcal{M}}(z, \zeta) = \frac{1}{(1 - z\bar{\zeta})^n} - \frac{(1 - |\alpha|^2)^n}{(1 - z\bar{\alpha})^n (1 - \alpha\bar{\zeta})^n}.$$

**The first approach:** Substituting (7.42) and (7.43) into (7.12) gives

$$B_j = (1 - |\alpha|^2)^{\frac{j+1-n}{2}}, \quad D_j = -\alpha \quad \text{for } j = 1, \dots, n.$$

We now get from (7.15)

$$F_n(z) = -\alpha + z \frac{1 - |\alpha|^2}{1 - z\bar{\alpha}} = b_{\alpha}(z),$$

and we get from (7.18)

$$F_{\ell}(z) = \frac{(z - \alpha)(1 - |\alpha|^2)^{\frac{n-\ell}{2}}}{(1 - z\bar{\alpha})^{n+1-\ell}} = b_{\alpha}(z) \cdot \left( \frac{\sqrt{1 - |\alpha|^2}}{1 - z\bar{\alpha}} \right)^{n-\ell} \quad \text{for } \ell = 1, \dots, n-1.$$

The identity (7.8) now amounts to

$$\begin{aligned} k_{\mathcal{M}}(z, \zeta) &= \frac{1}{(1 - z\bar{\zeta})^n} - \frac{(1 - |\alpha|^2)^n}{(1 - z\bar{\alpha})^n (1 - \alpha\bar{\zeta})^n} \\ &= \sum_{j=0}^{n-1} \frac{(1 - |\alpha|^2)^j b_{\alpha}(z) \overline{b_{\alpha}(\zeta)}}{(1 - z\bar{\alpha})^j (1 - \alpha\bar{\zeta})^j (1 - z\bar{\zeta})^{n-j}} \end{aligned}$$

and a related conclusion is that a function  $f$  belongs to  $\mathcal{M}$  if and only if it admits a representation

$$f(z) = b_{\alpha}(z) \cdot \sum_{j=0}^{n-1} \left( \frac{\sqrt{1 - |\alpha|^2}}{1 - z\bar{\alpha}} \right)^j \cdot g_{n-j}(z) \quad \text{for some } g_k \in \mathcal{A}_k. \quad (7.46)$$

**The second approach:** Since the function  $z \mapsto (1 - z\bar{\alpha})^{-j}$  belongs to  $H^\infty$  for all  $j$ , it follows that the factor  $\sum_{j=0}^{n-1} \left( \frac{\sqrt{1 - |\alpha|^2}}{1 - z\bar{\alpha}} \right)^j \cdot g_{n-j}(z)$  in (7.46) belongs to  $\mathcal{A}_n$ .

Therefore, the subspace  $\mathcal{M}$  can be represented as  $\mathcal{M} = b_{\alpha} \cdot \mathcal{A}_n$ ; this observation illustrates the second approach above. In this case,  $\mathcal{M}$  with lifted norm (7.21) is contractively included in  $\mathcal{A}_n$ .

**The third approach:** We first find  $B_k$  and  $D_k$  satisfying (6.22) for  $A$  and  $C$  as in (7.42) and hence, with  $\mathfrak{G}_{n,k,C,A}$  and  $\mathfrak{G}_{n,k+1,C,A}$  defined as in (7.44). We are

seeking  $B_k$  and  $D_k$  subject to

$$|B_k|^2 = \frac{1}{\mathfrak{G}_{n,k+1,C,A}} - \frac{|\alpha|^2}{\mathfrak{G}_{n,k,C,A}} = \frac{(1-|\alpha|^2)^n}{\mu_{n,k} \mathfrak{G}_{n,k,C,A} \mathfrak{G}_{n,k+1,C,A}}, \quad (7.47)$$

$$|D_k|^2 = \mu_{n,k} - \frac{(1-|\alpha|^2)^n}{\mathfrak{G}_{n,k,C,A}} = \mu_{n,k} \cdot \frac{|\alpha|^2 \mathfrak{G}_{n,k+1,C,A}}{\mathfrak{G}_{n,k,C,A}}, \quad (7.48)$$

$$B_k \overline{D}_k = -\frac{\overline{\alpha}(1-|\alpha|^2)^{\frac{n}{2}}}{\mathfrak{G}_{n,k,C,A}}. \quad (7.49)$$

Note that the second equalities in (7.48) and (7.49) follow from (7.44); alternatively one can use (7.45) combined with (2.7). The essentially unique choice of  $B_k$  and  $D_k$  satisfying (7.47)-(7.49) is the following:

$$B_k = -\frac{\overline{\alpha}}{|\alpha|} \cdot \frac{(1-|\alpha|^2)^{\frac{n}{2}}}{\sqrt{\mu_{n,k} \mathfrak{G}_{n,k} \mathfrak{G}_{n,k+1}}}, \quad D_k = |\alpha| \cdot \sqrt{\frac{\mu_{n,k} \mathfrak{G}_{n,k+1}}{\mathfrak{G}_{n,k}}}. \quad (7.50)$$

Now we arrive at the formula for  $\Theta_{n,k}$ :

$$\Theta_{n,k}(z) = \mu_{n,k}^{-1} D_k + z C R_{n,k+1}(zA) B_k \quad (7.51)$$

$$\begin{aligned} &= |\alpha| \cdot \sqrt{\frac{\mathfrak{G}_{n,k+1}}{\mu_{n,k} \mathfrak{G}_{n,k}}} - \frac{z \overline{\alpha} (1-|\alpha|^2)^n R_{n,k+1}(z \overline{\alpha})}{|\alpha| \sqrt{\mu_{n,k} \mathfrak{G}_{n,k} \mathfrak{G}_{n,k+1}}} \\ &= \frac{1}{|\alpha| \sqrt{\mu_{n,k} \mathfrak{G}_{n,k} \mathfrak{G}_{n,k+1}}} \cdot ((|\alpha|^2 \mathfrak{G}_{n,k+1} - (1-|\alpha|^2)^n z \overline{\alpha} R_{n,k+1}(z \overline{\alpha})) \\ &= \frac{(1-|\alpha|^2)^n}{|\alpha| \sqrt{\mu_{n,k} \mathfrak{G}_{n,k} \mathfrak{G}_{n,k+1}}} \cdot ((|\alpha|^2 R_{n,k+1}(|\alpha|^2) - z \overline{\alpha} R_{n,k+1}(z \overline{\alpha})) \\ &= \frac{(1-|\alpha|^2)^n}{|\alpha| \sqrt{\mu_{n,k} \mathfrak{G}_{n,k} \mathfrak{G}_{n,k+1}}} \cdot (R_{n,k}(|\alpha|^2) - R_{n,k}(z \overline{\alpha})), \end{aligned} \quad (7.52)$$

where we used (7.45) for the fourth equality and (2.7) for the fifth. For  $k=0$ , the latter formula gives

$$\begin{aligned} \Theta_{n,0}(z) &= \frac{(1-|\alpha|^2)^n}{|\alpha| \sqrt{\mu_{n,0} \mathfrak{G}_{n,0} \mathfrak{G}_{n,1}}} \cdot ((1-|\alpha|^2)^{-n} - (1-z \overline{\alpha})^{-n}) \\ &= \frac{1}{|\alpha| \sqrt{\mathfrak{G}_{n,1}}} \left( 1 - \left( \frac{1-|\alpha|^2}{1-z \overline{\alpha}} \right)^n \right) \\ &= -\frac{\overline{\alpha} b_\alpha(z)}{\sqrt{1-(1-|\alpha|^2)^n}} \cdot \sum_{j=0}^{n-1} \left( \frac{1-|\alpha|^2}{1-z \overline{\alpha}} \right)^j. \end{aligned} \quad (7.53)$$

For  $k \geq 1$ , we use the identity

$$\frac{1}{(1-|\alpha|^2)^j} - \frac{1}{(1-z \overline{\alpha})^j} = -\overline{\alpha} b_\alpha(z) \cdot \sum_{r=1}^j \frac{1}{(1-|\alpha|^2)^r (1-z \overline{\alpha})^{j-r}}$$

and formula (2.5) to get

$$\begin{aligned}
R_{n,k}(|\alpha|^2) - R_{n,k}(z\bar{\alpha}) &= \sum_{\ell=1}^n \binom{\ell+k-2}{\ell-1} \left[ \frac{1}{(1-|\alpha|^2)^{n+1-\ell}} - \frac{1}{(1-z\bar{\alpha})^{n+1-\ell}} \right] \\
&= -\bar{\alpha}b_\alpha(z) \sum_{\ell=1}^n \binom{\ell+k-2}{\ell-1} \sum_{r=1}^{n+1-\ell} \frac{1}{(1-|\alpha|^2)^r (1-z\bar{\alpha})^{n+1-\ell-r}} \\
&= -\bar{\alpha}b_\alpha(z) \sum_{\ell=1}^n \binom{\ell+k-2}{\ell-1} \sum_{j=0}^{n-\ell} \frac{1}{(1-|\alpha|^2)^{n+1-j\ell} (1-z\bar{\alpha})^j} \\
&= -\bar{\alpha}b_\alpha(z) \sum_{j=0}^{n-1} \left( \sum_{\ell=1}^{n-j} \binom{\ell+k-2}{\ell-1} \frac{1}{(1-|\alpha|^2)^{n+1-j-\ell}} \right) \cdot \frac{1}{(1-z\bar{\alpha})^j} \\
&= -\frac{\bar{\alpha}b_\alpha(z)}{(1-|\alpha|^2)^n} \cdot \sum_{j=0}^{n-1} \left( \sum_{\ell=0}^{n-j-1} \binom{\ell+k-1}{\ell} (1-|\alpha|^2)^\ell \right) \cdot \left( \frac{1-|\alpha|^2}{1-z\bar{\alpha}} \right)^j,
\end{aligned}$$

which being substituted into (7.52) gives

$$\Theta_{n,k}(z) = -\frac{\bar{\alpha}b_\alpha(z)}{|\alpha| \sqrt{\mu_{n,k} \mathfrak{G}_{n,k} \mathfrak{G}_{n,k+1}}} \cdot \sum_{j=0}^{n-1} \left( \sum_{\ell=0}^{n-j-1} \binom{\ell+k-1}{\ell} (1-|\alpha|^2)^\ell \right) \cdot \left( \frac{1-|\alpha|^2}{1-z\bar{\alpha}} \right)^j.$$

**The fourth approach:** Construction of the Bergman inner function associated with the wandering subspace for  $\mathcal{M}$  amounts to the construction of  $\Theta_0$  already derived above in (7.53). It is easily checked that the formula (7.53) agrees with [15, formula (6) page 125] (with  $n = 2$  in (7.53)), and with the formula for  $G_a(z)$  in [19, page 58] (with  $\alpha = n - 2$ ), up to a constant multiplicative factor. The derivation in both [15] and [19] is based on the work of Hedenmalm [20] whereby Bergman inner functions are produced as solutions of an appropriate extremal problem. We plan to discuss how our state-space methods can be used to solve such extremal problems directly in a future publication.

## 8. CONNECTIONS WITH TIME-VARYING LINEAR SYSTEMS THEORY

Here we make more explicit the connections of a Bergman-inner family  $\{\Theta_{n,k}\}_{k \geq 0}$  and the associated realization formulas (see Step 3 in the algorithm in Theorem 7.8) with the theory of conservative/dissipative time-varying linear systems as presented e.g. in [3] and [16].

We suppose that we are given an exactly  $n$ -observable pair  $(C, A)$  ( $A \in \mathcal{L}(\mathcal{X})$ ,  $C \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ ). Then Proposition 5.7 assures us that all the gramians  $\mathfrak{G}_{n,k,C,A}$  ( $k = 0, 1, 2, \dots$ ) are strictly positive definite. Let us introduce Hilbert spaces  $\mathcal{X}_k$  ( $k = 0, 1, 2, \dots$ ) with  $\mathcal{X}_k = \mathcal{X}$  for all  $k$  but with  $\mathcal{X}_k$  given the inner product induced by the  $k$ -th shifted gramian:

$$\langle x, x' \rangle_{\mathcal{X}_k} := \langle \mathfrak{G}_{n,k,C,A} x, x' \rangle_{\mathcal{X}}.$$

Similarly we set  $\mathcal{Y}_k = \mathcal{Y}$  with  $\mathcal{Y}_k$ -inner product given by

$$\langle y, y' \rangle_{\mathcal{Y}_k} := \left\langle \binom{n+k-1}{k} y, y' \right\rangle_{\mathcal{Y}}.$$

In addition suppose that operators  $B_k \in \mathcal{L}(\mathcal{U}_k, \mathcal{X})$  and  $D_k \in \mathcal{L}(\mathcal{U}_k, \mathcal{Y})$  are constructed as in Step 2 of the algorithm in Theorem 7.8. We then let  $\mathbf{U}_k$  be the colligation matrix

$$\begin{aligned}\mathbf{U}_k &= \begin{bmatrix} \mathbf{A}_k & \mathbf{B}_k \\ \mathbf{C}_k & \mathbf{D}_k \end{bmatrix} := \begin{bmatrix} \frac{k+n}{k+1}A & \binom{k+n}{k+1}B_k \\ C & \binom{k+n-1}{k}D_k \end{bmatrix} \\ &= \begin{bmatrix} \frac{k+n}{k+1}I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B_k \\ C & D_k \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \binom{k+n-1}{k}I_{\mathcal{U}_k} \end{bmatrix} : \begin{bmatrix} \mathcal{X}_k \\ \mathcal{U}_k \end{bmatrix} \rightarrow \begin{bmatrix} \mathcal{X}_{k+1} \\ \mathcal{Y} \end{bmatrix}. \quad (8.1)\end{aligned}$$

Here  $\mathbf{A}_k = \frac{k+n}{k+1}A$  with  $A$  considered as acting from  $\mathcal{X}_k$  into  $\mathcal{X}_{k+1}$  and similarly  $C_k = C$  but considered as acting from  $\mathcal{X}_k$  into  $\mathcal{Y}$  while  $B_k$  is considered as acting from  $\mathcal{U}_k$  into  $\mathcal{X}_k$ . The result of the construction in Theorem 7.8 is that each  $U_k := \begin{bmatrix} A & B_k \\ C & D_k \end{bmatrix}$  is unitary from  $\mathcal{X}_k \oplus \mathcal{U}_k$  to  $\mathcal{X}_{k+1} \oplus \mathcal{Y}_k$ . One can then study the associated conservative discrete-time time-varying linear system

$$\Sigma_{\{\mathbf{U}_k\}_{k \geq 0}} : \begin{cases} x(k+1) &= A_k x(k) + B_k u(k) \\ y(k) &= C_k x(k) + D_k u(k). \end{cases} \quad (8.2)$$

If we specify an initial condition  $x(0) = 0$ , then the input string  $\{u(k)\}_{k \geq 0}$  recursively uniquely determined via the system equations (8.2) an output string  $\{y(k)\}_{k \geq 0}$ , thereby defining a linear operator  $T_{\{\mathbf{U}_k\}} : \{u(k)\}_{k \geq 0} \mapsto \{y(k)\}_{k \geq 0}$ , the *input-output operator* for the linear system  $\Sigma_{\{\mathbf{U}_k\}_{k \geq 0}}$ . Such systems (with  $k$  running over all of  $\mathbb{Z}$  rather than just  $\mathbb{Z}_+$ ) are studied in [3] from the point of view of model theory for a sequence of contraction operators  $T_k = A_k^* : \mathcal{X}_{k+1} \rightarrow \mathcal{X}_k$  and associated time-varying unitary dilation theory and Lax-Phillips scattering theory. However the time-varying system associated with the realization of our Bergman-inner family of transfer functions is not given by  $\{U_k\}$  but rather by  $\{\mathbf{U}_k\}_{k \geq 0}$ , namely:

$$\Sigma_{\{\mathbf{U}_k\}_{k \geq 0}} : \begin{cases} x(k+1) &= \mathbf{A}_k x(k) + \mathbf{B}_k u(k) \\ y(k) &= \mathbf{C}_k x(k) + \mathbf{D}_k u(k). \end{cases} \quad (8.3)$$

where  $\mathbf{U}_k = \begin{bmatrix} \mathbf{A}_k & \mathbf{B}_k \\ \mathbf{C}_k & \mathbf{D}_k \end{bmatrix}$  is the tweak of  $U_k$  given by (8.1). The input-output operator  $T_{\{\mathbf{U}_k\}_{k \geq 0}}$  associated with the system (8.3) gives rise to an input-output operator  $T_{\{\mathbf{U}_k\}_{k \geq 0}} : \ell^2(\{\mathcal{U}_k\}_{k \geq 0}) \rightarrow \ell^2(\mathcal{Y})$  with image space  $T_{\{\mathbf{U}_k\}_{k \geq 0}}(\ell^2(\{\mathcal{U}_k\}_{k \geq 0}))$  exactly equal the set of all Taylor-coefficient strings  $\{y(k)\}_{k \geq 0}$  for which the associated  $Z$ -transform  $y(z) = \sum_{k=0}^{\infty} y(k)z^k$  is in the prescribed closed  $S_n$ -invariant subspace  $\mathcal{M}$  of  $\mathcal{A}_n(\mathcal{Y})$  (assuming that  $\mathcal{M}$  is chosen so that it is generated by its wandering subspace). Moreover  $T_{\{\mathbf{U}_k\}_{k \geq 0}}$  is isometric if the output string  $\{y(k)\}_{k \geq 0}$  is given the Bergman-space norm of its  $Z$ -transform  $\|\sum_{k=0}^{\infty} y(k)z^k\|_{\mathcal{A}_n(\mathcal{Y})}$ . It is amusing that this tweak  $U_k \mapsto \mathbf{U}_k$  as in (8.1) of a standard conservative time-varying linear system  $\Sigma_{\{U_k\}_{k \geq 0}}$  is the precise object required to get a realization theory for a Bergman-inner family  $\{\Theta_k\}_{k \geq 0}$ .

## REFERENCES

- [1] J. Agler, *Hypercontractions and subnormality*, J. Operator Theory **13** (1985), no. 2, 203–217.
- [2] A. Aleman, S. Richter and C. Sundberg, *Beurling's theorem for the Bergman space*, Acta Math. **177** (1996), no. 2, 275310.
- [3] D. Alpay, J.A. Ball, and Y. Peretz, *System theory, operator models and scattering: the time-varying case*, J. Operator Theory **47** (2002) 1 245–286.
- [4] C.-G. Ambrozie, M. Engliš, and V. Müller, *Operator tuples and analytic models over general domains in  $\mathbb{C}^n$* , J. Operator Theory **47** (2002), 287–302.

- [5] C. Apostol, H. Bercovici, C. Foias, and C. Pearcy, *Invariant subspaces, dilation theory, and the structure of the predual of a dual algebra, I*, J. Functl. Anal. **63** (1985), 369–404.
- [6] W. Arveson, *Subalgebras of  $C^*$ -algebras, III. Multivariable operator theory*, Acta Math. **181** (1998), 159–228.
- [7] J.A. Ball, V. Bolotnikov, and Q. Fang, *Multivariable backward-shift-invariant subspaces and observability operators*, Multidimens. Syst. Signal Process. **18** (2007) no. 4, 191–248.
- [8] J.A. Ball, V. Bolotnikov, and Q. Fang, *Transfer-function realization for multipliers of the Arveson space*, J. Math. Anal. Appl. **333** (2007) no. 1, 68–92.
- [9] J.A. Ball and V. Vinnikov, *Formal reproducing kernel Hilbert spaces: the commutative and noncommutative settings*, in *Reproducing Kernel Spaces and Applications* (Ed. D. Alpay), pp. 77–134, **OT 143**, Birkhäuser, Basel, 2003.
- [10] J.A. Ball and N. Cohen, *de Branges-Rovnyak operator models and systems theory: a survey*, in Topics in Matrix and Operator Theory, Rotterdam, 1989, (Ed. H. Bart, I. Gohberg, and M.A. Kaashoek), pp. 93–136, **OT 50** Birkhäuser-Verlag, 1991.
- [11] A. Beurling, *On two problems concerning linear transformations in Hilbert space*, Acta Math. **81** (1949), 239–255.
- [12] L. de Branges and J. Rovnyak, Canonical models in quantum scattering theory, in: *Perturbation Theory and its Applications in Quantum Mechanics* (C. Wilcox, ed.) pp. 295–392, Holt, Rinehart and Winston, New York, 1966.
- [13] L. de Branges and J. Rovnyak, *Square summable power series*, Holt, Rinehart and Winston, New York, 1966.
- [14] S. Chavan and R.E. Curto, *Operators Cauchy dual to 2-hyperexpansive operators: the multivariable case*, Integral Equations and Operator Theory **73** (2012), 481–516.
- [15] P. Duren and A. Shuster, *Bergman Spaces*, Mathematical Surveys and Monographs **100**, American Mathematical Society, 2004.
- [16] C. Foias, A.E. Frazho, I. Gohberg, and M.A. Kaashoek, *Metric Constrained Interpolation, Commutant Lifting and Systems*, **OT 100** Birkhäuser-Verlag, Basel, 1998.
- [17] I. Gohberg and L.A. Sakhnovich (Ed.), *Matrix and operator valued functions: the Vladimir Petrovich Potapov memorial volume*, **OT 72** Birkhäuser-Verlag, Basel, 1994.
- [18] P.R. Halmos, *Shifts on Hilbert spaces*, J. Reine Angew. Math. **208** (1961), 102–112.
- [19] H. Hedenmalm, B. Korenblum, and K. Zhu, *Theory of Bergman Spaces*, Graduate Texts in Mathematics **199**, Springer, 2000.
- [20] H. Hedenmalm, A factorization theorem for square area-integrable analytic functions, J. Reine Angew. Math. 422 (1991), 45–68.
- [21] C. Horowitz, *Factorization theorems for functions in the Bergman spaces*, Duke Math. J. **44** (1977), 201–213.
- [22] P.D. Lax, *Translation invariant subspaces*, Acta Math. **101** (1959), 163–178.
- [23] N.A. Loehr, *Bijective Combinatorics*, Discrete Mathematics and its Applications (Boca Raton), Chapman & Hall/CRC Press, Boca Raton, FL, 2011.
- [24] S. McCullough and T.T. Trent, *Invariant subspaces and Nevanlinna-Pick kernels*, J. Funct. Anal. **178** (2000) no. 1, 226–249.
- [25] V. Müller, *Models for operators using weighted shifts*, J. Operator Theory **20** (1988), no. 1, 3–20.
- [26] V. Müller and F.H. Vasilescu, Standard models for some commuting multioperators, Proc. Amer. Math. Soc. **117** No. 4 (1993), 979–989.
- [27] A. Olofsson, *A characteristic operator function for the class of  $n$ -hypercontractions*, J. Funct. Anal. **236** (2006), 517–545.
- [28] A. Olofsson, *An operator-valued Berezin transform and the class of  $n$ -hypercontractions*, Integral Equations Operator Theory **58** (2007), no. 4, 503–549.
- [29] A. Olofsson *Operator-valued Bergman inner functions as transfer functions*, Algebra i Analiz **19** (2007), no. 4, 146–173.
- [30] M. Rosenblum and J. Rovnyak, *Hardy classes and operator theory*, Oxford Univ. Press, New York, 1985; corrected reprint: Dover Publications, Mineola, NY, 1997.
- [31] S. Shimorin, *Wold-type decompositions and wandering subspaces for operators close to isometries*, J. Reine Angew. Math. **531** (2001), 147189.
- [32] S. Shimorin, *On Beurling-type theorems in weighted  $\ell^2$  and Bergman spaces*, Proc. Amer. Math. Soc. **131** (2003), no. 6, 1777–1787.

- [33] D.J. Sutton, *Structure of Invariant Subspaces for Left-Invertible Operators on Hilbert Space*, Virginia Tech PhD dissertation, August, 2010.

DEPARTMENT OF MATHEMATICS, VIRGINIA TECH, BLACKSBURG, VA 24061-0123, USA  
*E-mail address:* joball@math.vt.edu

DEPARTMENT OF MATHEMATICS, THE COLLEGE OF WILLIAM AND MARY, WILLIAMSBURG VA 23187-8795, USA  
*E-mail address:* vladimir.math.wm.edu